INFORMATION REQUIREMENTS FOR SELF-REPRODUCING SYSTEMS IN LUNAR ROBOTIC COLONIES

Amor Menezes, Pierre Kabamba

Department of Aerospace Engineering
University of Michigan
Ann Arbor, Michigan 48109-2140

amenezes@umich.edu, kabamba@umich.edu

September 15, 2006

KEYWORDS

Artificial life, information, moon, probabilistic generation theory, replication, reproduction, robotic colony, self-reproducing system, von Neumann threshold.

ABSTRACT

This paper examines the information requirement of a self-reproducing system in a lunar setting. Prior work is extended to allow self-reproducing entities the ability to select required resources in a probabilistic manner, leading to the notion of probabilistic reproduction. The quantitative requirement for probabilistic production of a non-degenerate offspring, i.e., an offspring with the same reproductive capability as the progenitor, is derived. The existence of a von Neumann rank threshold below which degeneracy always occurs is demonstrated.

The parallels between generation and communication are studied in some detail, and illustrated with examples. Utilizing the established results of Information Theory, it is shown that the channel capacity of a probabilistic generation system is the von Neumann information threshold, and the maximum reproduction rate of the system is the von Neumann complexity threshold.

The paper also proposes a simplified model of a lunar robotic colony to which the theoretical results are applicable.

1 INTRODUCTION

Artificially-created self-reproducing machines have long captivated the imagination of mankind. Scientific research conducted to realize the dream of self-reproduction has shown much promise in recent years, with the potential to significantly impact such diverse areas as space colonization, bioengineering, evolutionary software and autonomous manufacturing. This paper investigates one of the feasibility requirements of self-reproducing entities, and is set within the context of a lunar robotic colony.

1.1 Motivation

Current phased approaches to Martian exploration see the development of an enduring robotic presence on the Moon in the next five years. An International Lunar Robotic Village that will precede the establishment of a permanent human outpost on the Moon has also been proposed for 2014. This colony of advanced robots from the various co-operating space agencies would be designed to go beyond tele-presence from Earth, by sharing facilities that take advantage of available lunar resources, conducting life support experiments, and building the infrastructure necessary for human-robot Moon and Mars exploration. Recent space-exploration roadmaps suggest that individual countries will deploy these advanced robots on an as-needed basis to expand the size of the colony.

How much more efficient then, would it be to have robots endowed with the capacity for self-reproduction. These machines would be able to utilize available resources on-site to enlarge their numbers when deemed necessary for a given task. Such technology is not dependent on either the launch capabilities or the fiscal constraints surrounding the multiple launches of robots required for the colony, and therefore may provide a highly cost-effective solution.

In general, robots require mass, energy, and information or knowledge in order to perform their assigned tasks. While a seed robot endowed with some start-up resources is needed to initiate a colony, it is not feasible to deploy the robot with all the mass and energy necessary for self-reproduction, nor is it expected that the first seed robot arrives with all the knowledge it requires for survival, since it can learn from its actions within the environment. Accordingly, there are bounds on the minimum requirements of information, mass, and energy for the seed of a self-reproducing system. This work focuses solely on the information that is a fundamental requirement for the seed.

1.2 Background

Before proceeding any further, we should first state what is meant by the following terms that will be used throughout the paper: reproduction, replication, self-reproduction, and self-replication. For a historical perspective of the first two terms, the reader is referred to Freitas' excellent discussion on the subject in [2]. We consider reproduction in biological systems to imply the capacity for genetic mutations and the potential for evolution. Thus from an information standpoint, reproduction involves a change to the DNA code during the generation of progeny. Likewise, we will take *reproduction* in an artificial generation system to imply a change in the information specifications of an offspring. We reserve the term *replication* for progeny that have identical information content to that of the progenitor. *Self-reproducing* and *self-replicating* will be used to refer to those entities that perform the information equivalent of asexual reproduction or mitosis, i.e., the entities can reproduce or replicate based on the information specifications of only one progenitor.

The field of self-reproduction owes much to the efforts of John von Neumann [7], whose work on the theory of automata in the 1940s and 1950s inspired extensive research into the simulation and implementation of cellular automata, computer programs, kinematic machines, molecular machines, and even robotic colonies. A detailed overview of the research activities in the field is presented in [2, 6].

In a series of lectures in the 1940s [7], John von Neumann specified four functional requirements for reproduction: 1) a description of the progenitor; 2) a set of self-reproduction instructions; 3) a transfer of that description from the progenitor to its offspring; and, 4) an input to the process to allow for variation. He also postulated the existence of a threshold of complexity below which any attempt at self-reproduction was doomed to degeneracy. However, he did not define either complexity or degeneracy, nor did he go on to compute the threshold's value. An extensive literature survey in [4] indicates that no one had published an evaluation of this threshold in the following 60 years. Recently, [3] developed a novel theory of generation that is able to compute this von Neumann threshold, with results that yield a necessary and sufficient condition for non-degeneracy in self-reproduction.

1.3 Contribution

This work extends the results of [3] by proposing a probabilistic version of Generation Theory. This theory is different from that in [1], where a probabilistic measure of self-replicability is computed by comparing the probability of a machine spontaneously appearing in the environment to the probability that a new machine would appear, given that one already existed.

The new theory allows us to explicitly quantify the fundamental information requirement for self-reproduction. This requirement is imposed when probabilistically selecting from available resources to ensure that non-degenerate offspring are produced with high probability. We go on to demonstrate how similar reproduction and communication are to each other, and provide illustrative examples. Using established results from Information Theory, we develop the channel capacity and rate of a reproduction process, and relate it to quantities that have been previously defined in [3]: the von Neumann thresholds of information and complexity respectively.

1.4 Paper Layout

Section 2 presents the tenets of Probabilistic Generation Theory, Section 3 documents the parallels that exist with Information Theory, and Section 4 details a plan for a simplified model of a self-reproducing lunar robotic colony.

2 A Probabilistic Theory of Generation

The theory advanced here formalizes self-reproduction by "machines," a term describing any entity that is capable of producing an offspring regardless of its physical nature. Thus a robot, a bacterium, or even a piece of software code is considered to be a machine in this theory if they can each produce another robot, bacterium or some lines of code respectively. These machines require resources to self-reproduce, and each resource is chosen with some prior probability. The selected resource is then manipulated by the parent machine via an embedded generation action to produce an outcome, which itself may or may not be a machine. Thus we can state the following:

Definition 1. A *Probabilistic Generation System* is a quintuple $\Gamma = (U, M, R, P, G)$, where

- U is a *universal set* that contains machines, resources and outcomes of attempts at self-reproduction;

- $M \subseteq U$ is a *set of machines* in the context described;
- $R \subseteq U$ is a *set of resources* that can be utilized for self-reproduction;
- P is a probability mass function (pmf) on R, that is, $R \rightarrow \mathbb{R}$ with $P[r] \in [0,1]$ and $\sum_i P[r_i] = 1$;
- $G : M \times R \rightarrow U$ is a generation function that maps a machine and a resource into an outcome in the universal set, and not necessarily in the set of machines.

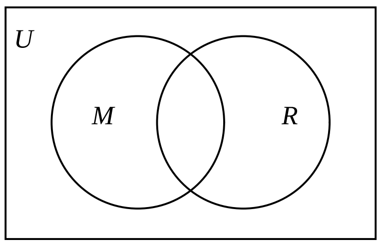


Figure 1: Pictorial representation of Definition 1.

Furthermore, it is possible that $M \cap R \neq \oslash$, and also $M \cup R \neq U$, as illustrated in Figure 1. The former implies that machines can belong to the set of resources, and the latter states that outcomes of attempts at generation may be neither machines nor resources.

One can define an indicator function, I, over a predicate, p, such that:

$$I(p) = 1 \; if \; p = True$$

$$I(p) = 0 \; if \; p = False.$$

Thus, the probability of a machine $x \in M$ processing a resource $r \in R$ to generate an outcome $y \in U$ may be written as:

$$P[y = G(x,r)] = \sum_{r \in R} I(y = G(x,r)) \cdot P[r]. \qquad (1)$$

If, in (1), $P[y = G(x,r)] > \varepsilon$, where $\varepsilon > 0$, then we say that "x is ε-capable of generating y," and we call the process *ε-reproduction*. If we have $P[x = G(x,r)] > \varepsilon$ in (1), where $\varepsilon > 0$, then we say that "x is ε-capable of generating itself," and we call the process *ε-replication*.

The four functional requirements for reproduction stipulated by von Neumann are captured above. There is a description (the parent, x), a set of self-reproduction instructions (the generation function, G), a transfer of that description (1), and an input to the process (the resource, r).

Of course, if we set $\varepsilon = 0$, then we allow every machine to ε-reproduce no matter what resource is selected. This is termed *Free Generation*. If $\varepsilon = 1$, then the deterministic theory of generation proposed in [3] is recovered, and only one resource is required to generate the desired offspring with probability 1. This is called *Strict Generation* or *Unity Generation*.

Definition 2. The *Generation Sets* in a probabilistic generation system are described as:

- $M_0 = M$, the set of all machines;
- M^{ε}_{i+1}, the set of all machines that are ε-capable of producing a machine of M^{ε}_i, $\forall i \geq 0$. That is, for $x \in M^{\varepsilon}_{i+1}$, $\exists\, y \in M^{\varepsilon}_i$ such that $P[y = G(x,r)] > \varepsilon$.

These sets are nested as indicated by the following proposition and Figure 2.

Proposition 1. $M_0 \supseteq M^{\varepsilon}_1 \supseteq M^{\varepsilon}_2 \supseteq \cdots \supseteq M^{\varepsilon}_i \supseteq M^{\varepsilon}_{i+1} \supseteq \cdots$.

Proof. See Appendix. □

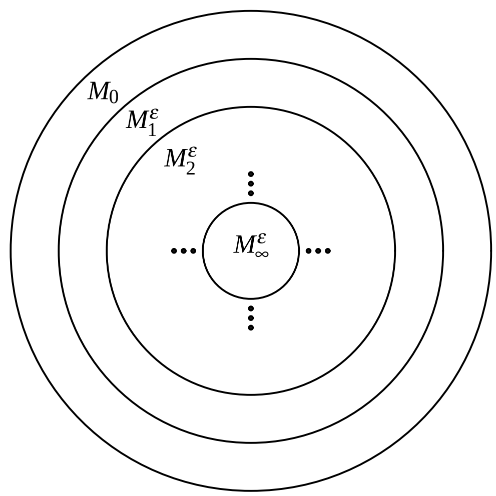


Figure 2: The nesting arrangement of the ε-generation sets.

As will be shown later in this section, the innermost generation set is important for self-reproduction. This set can be defined as:

$$M^{\varepsilon}_{\infty} = \bigcap_{i=0}^{\infty} M^{\varepsilon}_i. \qquad (2)$$

Proposition 2. *If $x \in M^{\varepsilon}_i \backslash M^{\varepsilon}_{i+1}$ and $P[y = G(x,r)] > \varepsilon$, $\varepsilon > 0$, then $y \notin M^{\varepsilon}_i$.*

Proof. See Appendix. □

The meaning of this proposition is that generation always proceeds outwards.

Corollary 1. *If $x \in M_i^\varepsilon \backslash M_{i+1}^\varepsilon$, then any sequence of ε-generation starting from x will produce an outcome $y \notin M$ in at most $i+1$ steps.*

Corollary 2. *If $x \in M_i^\varepsilon \backslash M_{i+1}^\varepsilon$, then it cannot ε-replicate.*

If replication is desired, Proposition 2 and its corollaries place a requirement on the location of a machine within the sets of a probabilistic generation sytem. To further ascertain the nature of this requirement, we develop the notion of rank, as well as a few other propositions.

Definition 3. The *rank of a probabilistic generation system*, $\rho^\varepsilon(\Gamma)$, where $\Gamma = (U,M,R,P,G)$ with generation sets M_i^ε, $i \geq 0$, is the smallest integer ρ such that $M_\rho^\varepsilon = M_{\rho+1}^\varepsilon$. If $\forall i, M_i^\varepsilon \neq M_{i+1}^\varepsilon$, then the generation system has infinite rank.

Proposition 3. *If $M_i^\varepsilon = M_{i+1}^\varepsilon$ for a probabilistic generation system with finite rank $\rho^\varepsilon(\Gamma)$, then $\forall j \geq i$, we have $M_j^\varepsilon = M_i^\varepsilon$.*

Proof. See Appendix. □

Hence, the nesting of the generation sets stop at the integer ρ for a probabilistic generation system of finite rank ρ. All generation sets of higher order (up to and including M_∞^ε) are equal. The next proposition indicates that if a probabilistic generation system has a finite number of machines, then its rank will always be finite.

Proposition 4. *For a probabilistic generation system $\Gamma = (U,M,R,P,G)$ where $\rho^\varepsilon(\Gamma)$ is finite, a finite number of machines, $|M|$, results in $\rho^\varepsilon(\Gamma) \leq |M|$.*

Proof. See Appendix. □

If $\rho^\varepsilon(\Gamma) = \infty$, then this implies that $|M| = \infty$. However, $|M| = \infty$ does *not* necessarily imply $\rho^\varepsilon(\Gamma) = \infty$. It is possible in the latter case for $\rho^\varepsilon(\Gamma)$ to be finite, including zero.

Having defined the rank of a probabilistic generation system, we can now go on to explain the rank of a machine.

Definition 4. The *rank of a machine*, $\rho^\varepsilon(x)$, in a probabilistic generation system $\Gamma = (U,M,R,P,G)$ with generation sets M_i^ε, $i \geq 0$, and $\rho^\varepsilon(\Gamma) = \rho$, is equal to i if $x \in M_i^\varepsilon \backslash M_{i+1}^\varepsilon$ ("deficient generation rank"), or is equal to ρ if $x \in \bigcap_{i=0}^\infty M_i^\varepsilon$ ("full generation rank").

This definition facilitates a key discussion on the generation-set location of a parent machine capable of ε-replication. We first consider the more general case of ε-generation cycles, defined as follows.

Definition 5. An *ε-generation cycle* is a sequence of ε-generations resulting in the production of a machine identical to itself after n generations.

Proposition 5. *If $x_1, x_2, \ldots, x_n$ form an ε-generation cycle of order n, then $x_i \in M_\infty^\varepsilon$, where $1 \leq i \leq n$.*

Proof. See Appendix. □

Corollary 3. *If a machine, x, is capable of ε-replication (an ε-generation cycle of order one) in a probabilistic generation system, then $x \in M_\infty^\varepsilon$.*

We now show that the *Principle of Degeneracy* stated in [3] also holds true in the probabilistic version of the theory. Having identified that machines capable of ε-replication must belong to M_∞^ε, and that any exit from M_∞^ε is irreversible, we demonstrate that it is possible for offspring machines to belong to M_∞^ε as long as their parents do as well.

Proposition 6. *Assume that a probabilistic generation system $\Gamma = (U,M,R,P,G)$ has finite rank $\rho^\varepsilon(\Gamma) = \rho$, and let $x \in M_\infty^\varepsilon$. Then $\exists\, y \in M_\infty^\varepsilon$ such that $P[y = G(x,r)] > \varepsilon$.*

Proof. See Appendix. □

Corollary 4. *Assume that a probabilistic generation system $\Gamma = (U,M,R,P,G)$ has finite rank $\rho^\varepsilon(\Gamma) = \rho$, and also has $|M_\infty^\varepsilon| \neq 0$ and $|M_\infty^\varepsilon| < \infty$. Then M_∞^ε contains at least one ε-generation cycle of order at most $|M_\infty^\varepsilon|$.*

Proposition 6 quantifies non-degenerate ε-reproduction and ε-replication. It corroborates von Neumann's remarks, and indicates that there is a minimum threshold beyond which a machine is able to ε-generate an offspring without a decrease in generation rank. We call this the *von Neumann Rank Threshold*, τ_r, and define

$$\tau_r^\varepsilon = \rho^\varepsilon(\Gamma). \tag{3}$$

We can also describe an interesting fact about the minimum number of resource elements in a probabilistic generation system, given certain assumptions.

Proposition 7. *Assume that a probabilistic generation system $\Gamma = (U,M,R,P,G)$ has finite rank $\rho^\varepsilon(\Gamma) = \rho$, and that a machine $y_1 \in M_{\rho-1}^\varepsilon \backslash M_\rho^\varepsilon$ can be ε-generated, where $0.5 < \varepsilon \leq 1$. Then the resource set, R, must contain at least two elements.*

Proof. See Appendix. □

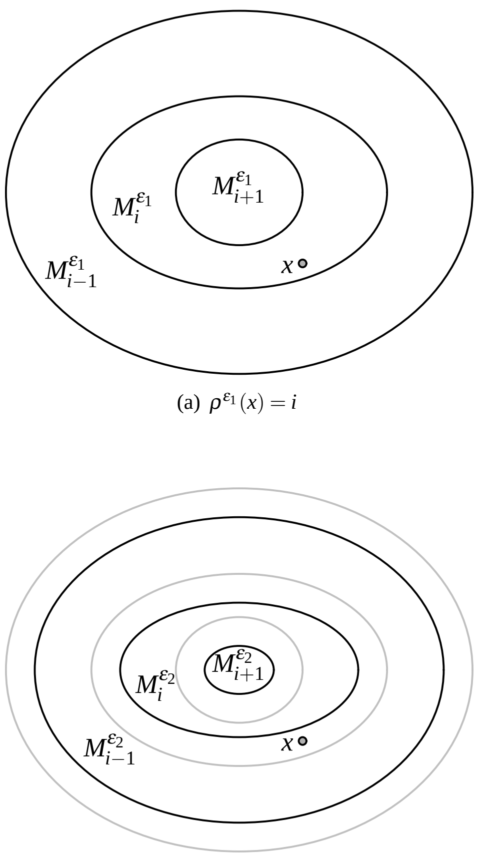
Corollary 5. *If every machine can be ε-generated in a probabilistic generation system $\Gamma = (U,M,R,P,G)$, where $0.5 < \varepsilon \leq 1$, and the resource set is a singleton, then the rank of the system, $\rho^\varepsilon(\Gamma)$, must be either 0 or ∞.*

We now proceed to analyze the effects of changes in ε on the generation sets of the system.

Proposition 8. *Given a probabilistic generation system* $\Gamma = (U, M, R, P, G)$*, if* $\varepsilon_1 < \varepsilon_2$*, then* $M_i^{\varepsilon_1} \supseteq M_i^{\varepsilon_2}$*, for all* $i \geq 0$.

Proof. See Appendix. □

Corollary 6. *If* $x \in M_i^{\varepsilon_1} \setminus M_{i+1}^{\varepsilon_1}$*, and* $\varepsilon_1 < \varepsilon_2$*, then* $\rho^{\varepsilon_1}(x) \geq \rho^{\varepsilon_2}(x)$.



(a) $\rho^{\varepsilon_1}(x) = i$

(b) $\rho^{\varepsilon_2}(x) = i - 1$

Figure 3: The effect of increased ε on machine rank.

Figure 3 illustrates Corollary 6.

It turns out that the rank of a machine is not the only thing affected by a change in ε. The rank of a probabilistic generation system, and hence the von Neumann threshold, also vary in the manner indicated below.

Proposition 9. *Given a probabilistic generation system* $\Gamma = (U, M, R, P, G)$*, if* $\varepsilon_1 < \varepsilon_2$*, then* $\rho^{\varepsilon_1}(\Gamma) \leq \rho^{\varepsilon_2}(\Gamma)$.

Proof. See Appendix. □

This proposition makes intuitive sense. If the generation sets become bigger with decreasing ε, then it stands to reason that more machines will be able to ε-replicate, and therefore, $\tau_r^\varepsilon = \rho^\varepsilon(\Gamma)$ must be correspondingly reduced. In fact, an even stronger statement can be made.

Proposition 10. *Given a probabilistic generation system where* $\Gamma = (U, M, R, P, G)$ *with finite rank* $\rho^\varepsilon(\Gamma) = \rho$, $\varepsilon \to 0$ *implies that* $M_\infty^\varepsilon \to M_0$.

Proof. See Appendix. □

Having formalized probabilistic self-reproduction with the above definitions and propositions, we now demonstrate the applicability of the theory and its usefulness as a tool for analysis.

3 The Information Theory Analogue

The processes of generation and communication [5] display remarkable parallels; indeed, communication may be viewed as the reproduction of a transmitted message at the receiving end of a communication channel. Table 1 indicates the full extent of these similarities.

To explain the parallels fully, consider the typical diagram of a communication process [5] modified in accordance with Table 1 as shown in Figure 4, in order to yield a corresponding diagram of a generation process.

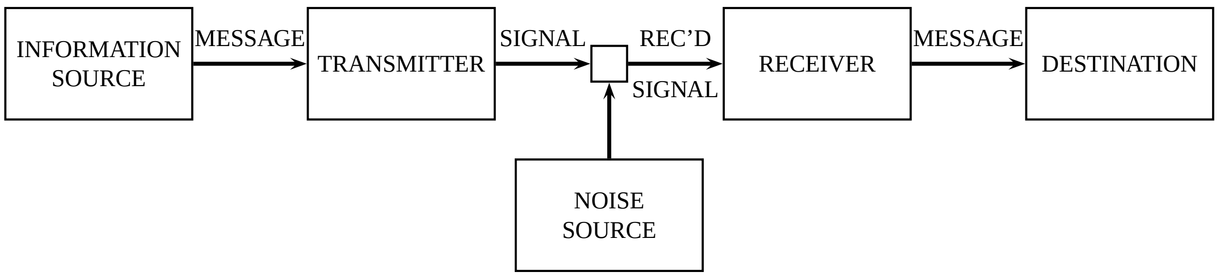
The parent machine acts as the information source, producing the instructions (a message) that will create an offspring. These instructions can be coded, just like DNA, prior to the generation process (correspondingly, the message can be encoded into a syntactically correct form prior to transmission). There is a non-zero probability that mutations may occur during reproduction (the message may be corrupted by noise), and different resources (various noise samples) will produce different outcomes. If the probability of producing a certain outcome exceeds a threshold, then that outcome is produced.

Typically when we write $y = G(x, r)$, we combine the information source and the transmitter into x, and the receiver and destination into y, as depicted in Figure 4.

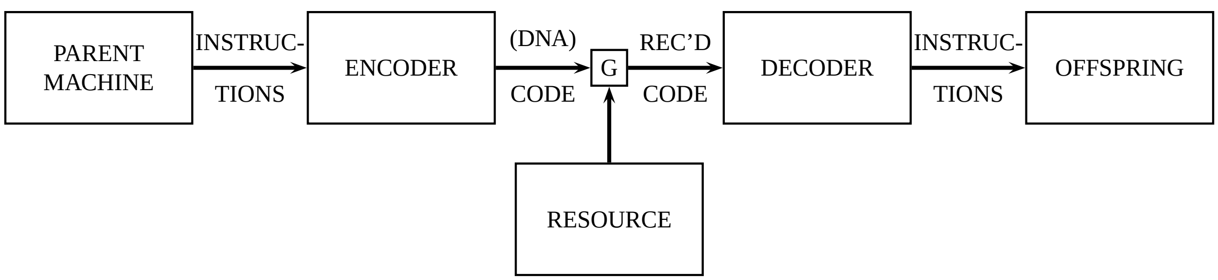
The next example depicts how probabilistic generation theory may be successfully utilized in a binary communication channel, and also sheds some light on the changes in the generation sets as a result of varying ε.

Example 1. *Binary Communication Channel With Probabilistically Selected Noise Sequences*

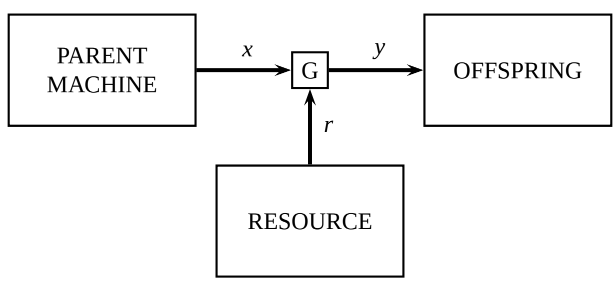
Let M, the set of messages, be represented by a set of syntactically correct binary sequences, where the syntax is as follows: a length of 4 symbols; and the number of ones in the sequence is even and greater than zero.



(a) Typical communication system [5].



(b) Expanded generation system.



(c) Typical generation system.

Figure 4: Schematic diagrams of communication and generation systems.

Table 1: Comparison of the Probabilistic Generation System in Generation and Information Theories

	Generation Theory	Information Theory
U	Universal set of machines, resources and outcomes of attempts at self-reproduction	Universal set of representations of messages, representations of noise samples, and outcomes of attempts at communication
M	Set of machines	Set of syntactically correct representations of messages
R	Set of resources	Set of representations of noise samples that, when used in a communication channel, deterministically alter a portion of the transmitted message
P	pmf on R (probability of selection)	pmf on R (probability of selection)
G	Generation function, $G: M \times R \to U$	communication channel where the altering of a message occurs deterministically based on the probabilistic choice of noise sample, $G: M \times R \to U$
$M_0 \backslash M_1^\varepsilon$	Set of machines that, when attempting to ε-reproduce, create offspring that are never machines no matter which resource is selected	Set of syntactically correct messages that, when transmitted, are interpreted as messages that are never syntactically correct no matter which noise sample is selected
$M_i^\varepsilon \backslash M_{i+1}^\varepsilon$	Set of machines such that, for all sequences of $i+1$ resources, these machines ε-produce an offspring that is not a machine at some generation in the sequence	Set of syntactically correct messages such that, for all sequences of $i+1$ noise samples, the recursive transmission of these messages results in an outcome that can never be interpreted as a syntactically correct message at some communication in the sequence
M_∞^ε	Set of machines such that there exists an infinitely long sequence of resources for which all the ε-offsprings are machines	Set of syntactically correct messages such that there exists an infinitely long sequence of noise samples for which all the transmitted messages are syntactically correct
ε	$\varepsilon \in [0,1]$	$\varepsilon = 0$

Thus the set of acceptable messages is

$$\{1111, 1100, 1010, 1001, 0110, 0101, 0011\}.$$

Let R, the set of noise sequences, be represented by a set of binary sequences with the following characteristics: a length of 4 symbols; either the first bit or the second bit is a one; and either the third bit or the fourth bit is a one.

Thus the set of noise sequences is

$$\{1001, 0101, 1010, 0110\}.$$

Let P be a pmf on R, and taken for this example to be

$$\{0.5, 0.1, 0.15, 0.25\}.$$

Let G be a binary communication channel as indicated in Figure 5. This channel has crossover probabilities given by the second and third bits of the selected noise resource, and correct-transmission probabilities given by the first and fourth bits of the selected noise resource.

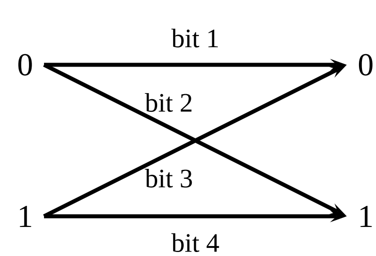


Figure 5: Binary communication channel representing G in Example 1.

Suppose that the message $x = 1111$ is given, and that the output message (or ε-offspring) desired is $y = 1111$, with $\varepsilon = 1$.

Consider $y_1 = G(x, r_1)$, where $r_1 = 1001$. As shown in the diagram of Figure 6, the crossover probabilities are both 0, and as a result, $y_1 = 1111$.

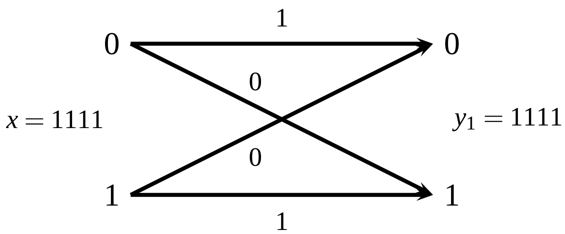


Figure 6: $y_1 = G(x, r_1)$.

Now consider $y_2 = G(x, r_2)$, where $r_2 = 0101$. In this case, the crossover probability is 1 if a 0 is transmitted, and a 0 if a 1 is transmitted. Since there are no zeroes in the original message, we still get $y_2 = 1111$. The diagram in Figure 7 makes this clear.

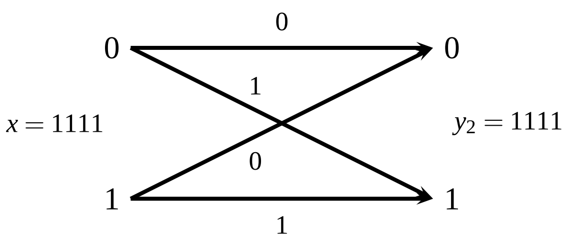


Figure 7: $y_2 = G(x, r_2)$.

With $y_3 = G(x, r_3)$, where $r_3 = 1010$, we have a crossover probability of 0 if a 0 is transmitted, but it is a 1 if a 1 is transmitted. Thus all the ones in the original message are transformed, and we get $y_3 = 0000$ as indicated in Figure 8. By the syntax defined earlier, this outcome is not considered to be a valid message.

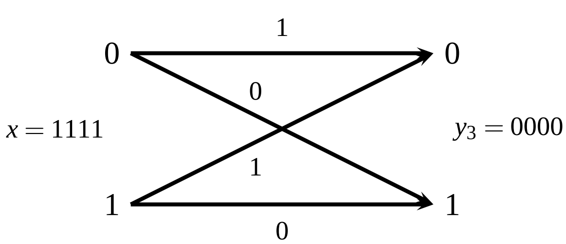


Figure 8: $y_3 = G(x, r_3)$.

Lastly, for $y_4 = G(x, r_4)$, where $r_4 = 0110$, we have a similar result. Here, the crossover probabilities are both 1, and as a result, both inputs of 0 and 1 are switched. We have $y_4 = 0000$ (Figure 9).

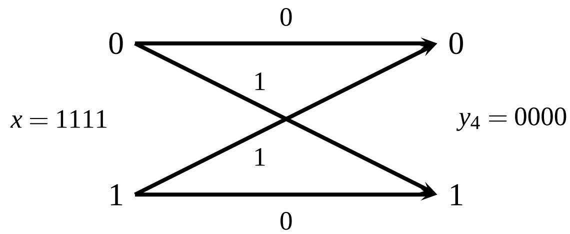


Figure 9: $y_4 = G(x, r_4)$.

Thus we can calculate $P[y = G(x, r)] = 1 \times 0.5 + 1 \times 0.1 + 0 \times 0.15 + 0 \times 0.25 = 0.6$.

Clearly, $P[y = G(x, r)] < 1$, and so ε-generation for the message $x = 1111$ will only occur if $\varepsilon < 0.6$.

This implies that $x = 1111$ belongs to the set $M_0 \backslash M_1^1$, with $\rho^1(\Gamma) = 1$. Any messages transmitted are disregarded since the threshold is set too high.

Also, $x = 1111$ belongs to the set $M_1^{0.55}$, and $\rho^{0.55}(\Gamma) = 1$. With the lower threshold, x belongs to the set of messages that can be perfectly transmitted. This is true for any ε that is less than 0.6.

As an aside, note that if we require the output message to be 1100 (or even any of the other possible messages: 1010,1001,0110,0101,0011), then $P[y = G(x, r)] = 0$, since these messages cannot be generated from the given message based on the noise sequences available.

We can carry out a similar analysis for each of the possible messages listed previously. These results are summarized in Table 2, along with the probability that the output message is the input message itself. The table helps fully explain the set structures illustrated in Figure 10, for ε values of 1, 0.55, and 0. Since $P[x = G(x, r)] = 0.6$ for $x = 1111$, and 0.5 for all other messages, we have that for any $\varepsilon \leq 0.5$, $M_0 = M_\infty^\varepsilon$, and so $\rho^\varepsilon(\Gamma) = 0$. The figure also illustrates the results of Propositions 8 and 9.

It is apparent that there is a noise sequence in the set of resources, $\{1010\}$, such that the transmitted output is never a syntactically correct message, i.e., a string of zeroes is always produced. Consequently, the set $M_0 \backslash M_1^\varepsilon$ is never empty in this example. Now in most communication systems employed today, the converse is true, since decoders serve to decode the received string into the closest possible message that is syntactically correct. As a result, $|M_0 \backslash M_1^\varepsilon|$ is reduced as much as possible.

Lastly, there is also a resource, $\{1001\}$, such that the transmitted output is always syntactically correct, and indeed, the original input itself. Hence if we impose the requirement that $P[x = G(x, r)]$ be high, i.e. $\varepsilon \approx 1$, then we are necessarily imposing a condition on the binary communication channel that it be close to identity. This condition can only be achieved with r_1. In fact, this result cor-

Table 2: Message Outputs for Various Noise Sequences

	$y_1 = G(x,r_1)$	$y_2 = G(x,r_2)$	$y_3 = G(x,r_3)$	$y_4 = G(x,r_4)$	$P[x = G(x,r)]$
$x = 1111$	1111	1111	0000	0000	0.6
$x = 1100$	1100	1111	0000	0011	0.5
$x = 1010$	1010	1111	0000	0101	0.5
$x = 1001$	1001	1111	0000	0110	0.5
$x = 0110$	0110	1111	0000	1001	0.5
$x = 0101$	0101	1111	0000	1010	0.5
$x = 0011$	0011	1111	0000	1100	0.5

roborates Corollary 5, because we now have the only possible singleton resource such that $\tau_r^1 = 0$. Thus there are only two ways to induce a rank of 0 for this example: 1) reduce ε, at the price of low-probability results; or 2) use a specific singleton resource to yield a high-probability outcome, but this is unrealistic because one cannot typically specify noise samples. This concludes the example.

Specifying a machine in M_∞^ε requires an amount of information τ_i^ε, and this information threshold is given by:

$$\tau_i^\varepsilon = log_2 \frac{|M_0|}{|M_\infty^\varepsilon|} \text{ for } |M_\infty^\varepsilon| \neq 0 \quad (4)$$

$$= \infty \text{ for } |M_\infty^\varepsilon| = 0 \quad (5)$$

for a particular ε. But what if ε-replication or an ε-generation cycle requires machines of lower-rank to be utilized as a resource? For instance, a look back at the previous example shows that acceptable messages and noise samples could both be specified by identical syntax, and indeed $R \subset M$. Thus there were messages that also doubled as possible noise samples too, and it is conceivable that machines can make use of other lower-rank machines in order to propagate. Then the question arises: how much information would be required to specify each of these lower-rank machines?

To answer, consider that we have demonstrated a strong likeness between Probabilistic Generation Theory and Information Theory. It would therefore serve to make use of established results in Information Theory, and identify a corresponding interpretation in ε-reproduction. This is best illustrated with another example.

Example 2. *A Strict Generation System*

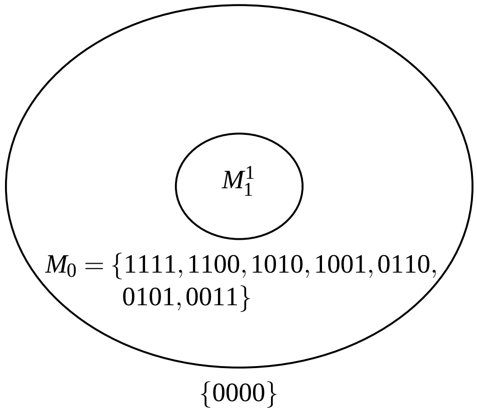
Suppose we are given the strict generation system in Figure 11, where we know the complete list of parents and offspring *a priori* as follows:

$$\begin{aligned}
x_2 &= G(x_1,r_1)\\
x_3 &= G(x_1,r_2)\\
x_4 &= G(x_1,r_3)\\
x_5 &= G(x_2,r_1)\\
x_6 &= G(x_2,r_2)\\
x_7 &= G(x_3,r_1)\\
x_8 &= G(x_4,r_1)\\
x_9 &= G(x_5,r_1)\\
x_{10} &= G(x_5,x_9)\\
x_{11} &= G(x_6,x_{10})\\
x_{12} &= G(x_6,x_{11})\\
x_{13} &= G(x_7,x_{12})\\
x_{14} &= G(x_7,x_{13})\\
x_{15} &= G(x_8,x_{14})\\
x_{16} &= G(x_8,x_{15})\\
x_1 &= G(x_1,x_{16}).
\end{aligned}$$

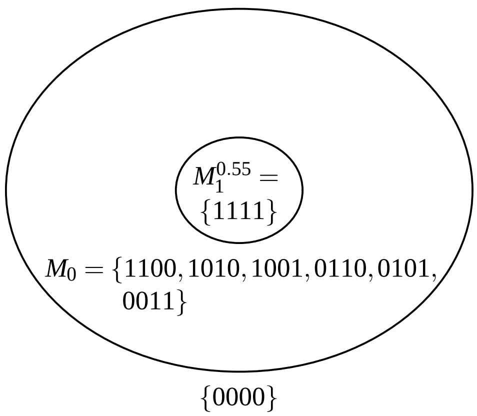
Utilizing the Generation Analysis Algorithm in [3], the generation sets are obtained as indicated in Figure 12, where:

$$\begin{aligned}
M_0 \backslash M_1^1 &= \{x_9,x_{10},x_{11},x_{12},x_{13},x_{14},x_{15},x_{16}\}\\
M_1^1 \backslash M_2^1 &= \{x_5,x_6,x_7,x_8\}\\
M_2^1 \backslash M_3^1 &= \{x_2,x_3,x_4\}\\
M_3^1 &= \{x_1\} \text{ and}\\
\rho^1(\Gamma) &= 3.
\end{aligned}$$

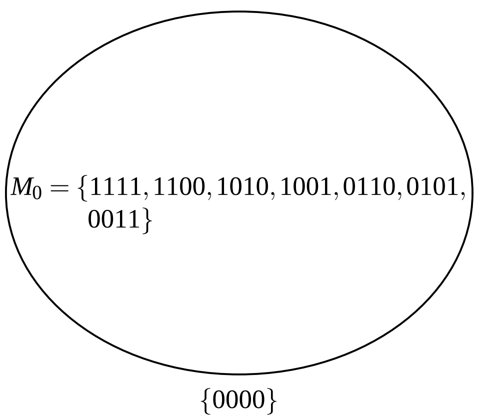
This contrived system is thus dependent on lower-rank machines for the replication of x_1. In accordance with the form of τ_i^ε, we can also write $\iota(x \in M_i^\varepsilon)$, the information



(a) $\varepsilon = 1$; $\rho^1(\Gamma) = 1$



(b) $\varepsilon = 0.55$; $\rho^{0.55}(\Gamma) = 1$



(c) $\varepsilon \leq 0.5$; $\rho^{\varepsilon}(\Gamma) = 0$

Figure 10: The ε-generation set structure of the system in Example 1.

of a machine belonging to M_i^{ε} as

$$\iota(x \in M_i^{\varepsilon}) = log_2 \frac{|M_0|}{|M_i^{\varepsilon}|}. \qquad (6)$$

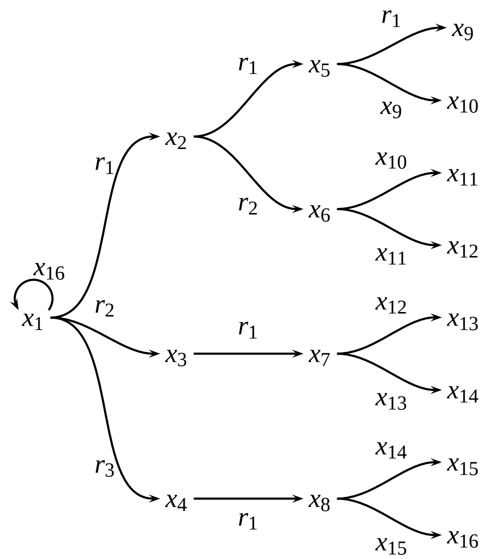


Figure 11: Generation Diagram for Example 2.

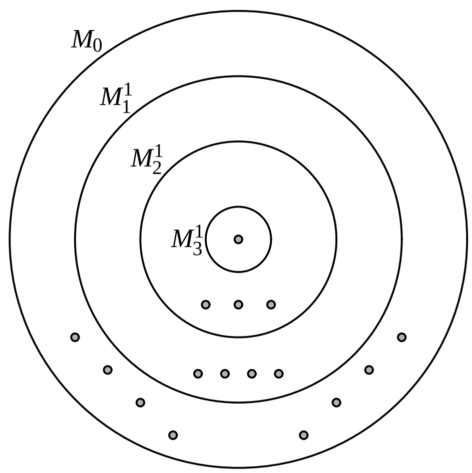


Figure 12: Generation set structure of Example 2.

We can calculate these quantities to be:

$$\begin{aligned}
\tau_i^1 &= log_2 \frac{|M_0|}{|M_\infty^1|} = log_2 16 = 4 \text{ bits}, \\
\iota(x \in M_2^1) &= log_2 \frac{|M_0|}{|M_2^1|} = 2 \text{ bits}, \\
\iota(x \in M_1^1) &= log_2 \frac{|M_0|}{|M_1^1|} = 1 \text{ bit}, \\
\iota(x \in M_0^1) &= log_2 \frac{|M_0|}{|M_0^1|} = 0 \text{ bits},
\end{aligned}$$

and so with a four bit code we have that:

- machines in M_0^1 have four free bits in their code;
- machines in M_1^1 have three free bits in their code,

e.g., their four bit code starts with a 1;

- machines in M_2^1 have two free bits in their code, e.g., their four bit code starts with a 11;
- machines in M_3^1 have no free bits in their code.

This implies that $x_1 = 1111$, and so the remaining bit combinations in M_2^1 are

$$\{1100, 1110, 1101\};$$

the remaining bit combinations in M_1^1 are

$$\{1000, 1001, 1010, 1011\};$$

and the remaining bit combinations in M_0 are those not already specified.

Upon closer examination, it is evident that $M_0 \backslash M_1^1$ specifies an extra bit (and there are only 3 free), since the bit combinations allowed have to start with a 0. Similarly, $M_1^1 \backslash M_2^1$ specifies an extra bit (with only 2 free), because the bit combinations allowed have to start with 10. $M_2^1 \backslash M_3^1$ specifies more than 2 bits, since the bit combinations have to start with 11, and in addition, the combination 1111 is disallowed. Let us introduce the notion of *Rank Information*, and define it as follows:

$$\iota(x \in M_i^\varepsilon \backslash M_{i+1}^\varepsilon) = -log_2 \frac{\left|M_i^\varepsilon \backslash M_{i+1}^\varepsilon\right|}{|M_0|}. \qquad (7)$$

Checking with the example to see what this gives us,

$$\begin{aligned}
\iota(x \in M_0 \backslash M_1^1) &= -log_2 \frac{8}{16} = 1 \text{ bit}, \\
\iota(x \in M_1^1 \backslash M_2^1) &= -log_2 \frac{4}{16} = 2 \text{ bits}, \\
\iota(x \in M_2^1 \backslash M_3^1) &= -log_2 \frac{3}{16} > 2 \text{ bits}, \\
\iota(x \in M_3^1) &= -log_2 \frac{1}{16} = 4 \text{ bits},
\end{aligned}$$

as desired.

We can generalize this example to any generation system that is specified completely. Now in $y = G(x, r)$, probabilistic r implies that y, and hence $\rho^\varepsilon(y)$ are also probabilistic. In fact, since we are aware of the entire generation system beforehand, we can calculate the probability of the rank of an offspring as:

$$\begin{aligned}
\pi_i^\varepsilon &= P[x \in M_i^\varepsilon \backslash M_{i+1}^\varepsilon] && (8) \\
&= \frac{\left|M_i^\varepsilon \backslash M_{i+1}^\varepsilon\right|}{|M_0|}, \text{ for } 0 \le i < \rho. && (9) \\
\pi_\rho^\varepsilon &= P[x \in M_\rho^\varepsilon] && (10) \\
&= \frac{|M_\infty^\varepsilon|}{|M_0|}. && (11)
\end{aligned}$$

Therefore, the Rank Information satisfies

$$\iota(x \in M_i^\varepsilon \backslash M_{i+1}^\varepsilon) = -log_2 \pi_i^\varepsilon,\ 0 \le i \le \rho. \qquad (12)$$

The *Rank Entropy* of the probabilistic generation system is:

$$H_r^\varepsilon = -\sum_{i=0}^{\rho} \pi_i^\varepsilon log_2 \pi_i^\varepsilon. \qquad (13)$$

Making use of Information Theory, the capacity of a memoryless channel is the maximum of the mutual information over all the possible input probabilities. By construction, the equivalent here amounts to taking the maximum of the Rank Information, with the result that the reproductive capacity of a probabilistic generation process is τ_i^ε, the von Neumann information threshold. This result makes it clear that there is only a certain amount of information required to completely specify a machine, and when this maximum is reached during self-reproduction, the von Neumann system rank is achieved.

We can also use Shannon's channel coding theorem, which states that the transfer rate of a code (in number of symbols per second) is less than or equal to the channel capacity divided by the entropy of the source. We have a similar notion here, i.e., the reproduction rate (in symbols of generation code per generation), η^ε, is constrained in the manner:

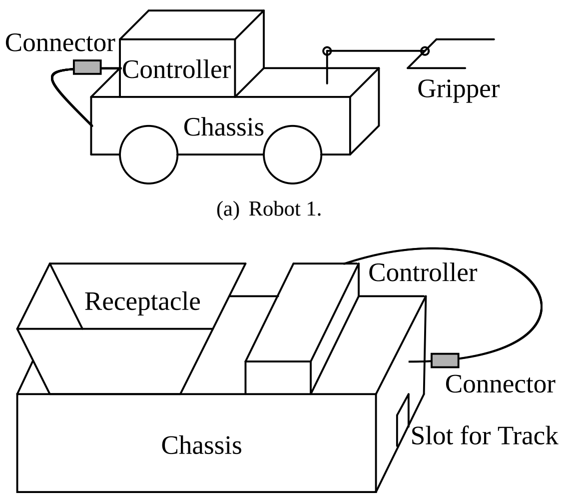
$$\eta^\varepsilon \le \frac{\tau_i^\varepsilon}{H_r^\varepsilon}. \qquad (14)$$

The quantity on the right in the above equation is the von Neumann complexity threshold, τ_c^ε generalizing that in [3]. Full transfer of complexity implies that the rate of self-reproduction is at a maximum, and the offspring has the maximum amount of information needed to achieve the von Neumann system rank.

4 Application to Lunar Robotic Colonies

In our efforts to develop self-reproduction as a cost-effective solution for the establishment of robotic outposts, we are planning a physical generation system that simulates the operation of a growing robot colony. It is anticipated that, in order to create the International Lunar Robotic Village, the advanced robots provided by the various co-operating space agencies will be multi-disciplinary in nature, with each robot capable of performing assorted tasks such as mining for lunar resources as well as building any necessary infrastructure. These robots will be structurally different from one another and, although their functions will necessarily be symbiotic, individuals will only make decisions for the good of the colony. Thus, self-reproduction will take place in order to benefit the group as whole.

Using off-the-shelf components, the planned generation system is a simplified version of a robotic village, initially consisting of two robots. The first robot has a controller on board a wheeled chassis, and is equipped with a gripper. The second robot also has a controller on board a chassis, but this robot slides along a track, has a storage receptacle, and does not have a gripper. Both robots are equipped with the necessary sensors to perceive the environment. Figure 13 depicts what these robots would look like.



(a) Robot 1.

(b) Robot 2.

Figure 13: Schematic representation of the robots in the generation system.

The objective of the colony is to transfer a number of materials from a location near the middle of the track to a location at one of the ends of the track. Each piece of material is gripped by Robot 1 and placed in the receptacle of Robot 2, which then transports it to the end of the track and drops it there (Figure 14). Rather than have the robots co-operatively determine whether it is advantageous to expand the colony and pursue self-reproduction, we specify that each robot self-reproduce after a period of time. This eliminates any decision-making complications.

Robot 1 self-reproduces by moving a short distance away from the work site to where a number of Robot 1-type controllers and chassis are located. It uses the gripper to connect a controller to its chassis, thereby activating a new Robot 1 (Figure 15(a)).

Robot 2 begins self-reproduction by moving to the end of the track opposite from the drop-off location. Here, there are a number of unconnected controllers on board Robot 2-type chassis. Robot 2 nudges the electrical connector on the controller when it arrives at this location, thereby activating a new Robot 2 (Figure 15(b)).

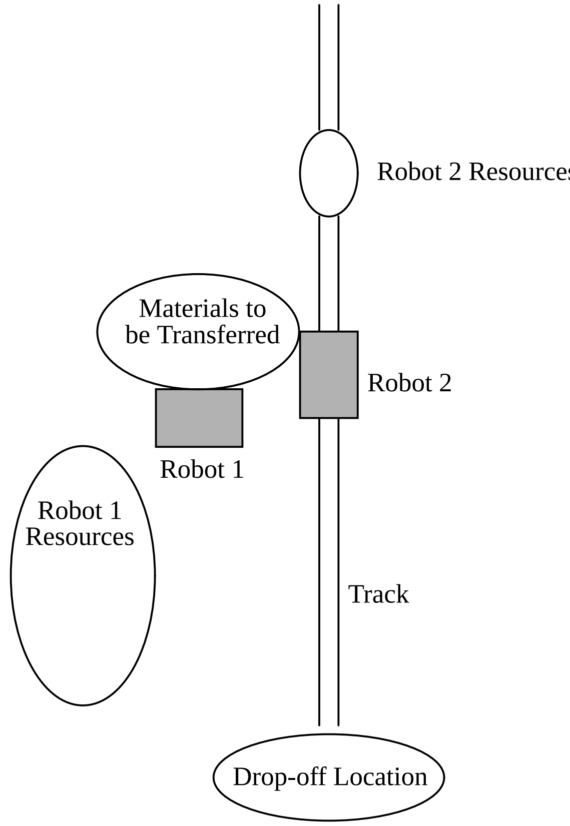
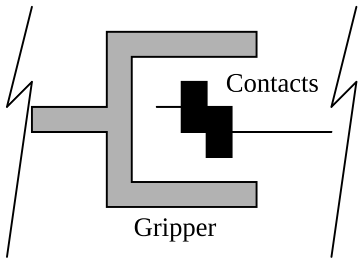


Figure 14: Proposed colony layout.

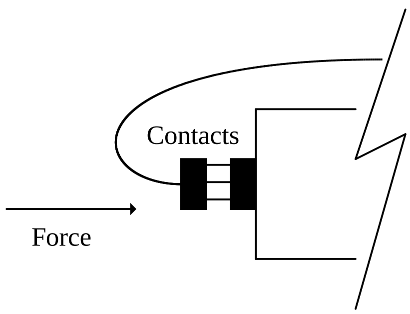
In this system,

$$M = \{x_1, x_2\}$$
$$R = \{r_1, r_2\},$$

where x_1 = Robot 1, x_2 = Robot 2, r_1 = unconnected Robot 1 controller-chassis combo, and r_2 = unconnected Robot 2 controller-chassis combo. The generation functions for each robot involve the manner in which the controller is connected to the chassis; we have $x_1 = G(x_1, r_1)$, and $x_2 = G(x_2, r_2)$. If x_1 operates on r_2 or if x_2 operates on r_1, then any attempt at self-reproduction will not produce a machine. This is because proper electrical connection of the desired robot will not have been achieved. For the generation system specified above, we are able to apply Probabilistic Generation Theory and evaluate the following parameters: the von Neumann thresholds of rank, information and complexity; channel capacity; reproduction rate; and the Rank Information and Rank Entropy of the generation sets. Full details and results on the above generation system will be presented in a future paper.



(a) Robot 1 squeezes contacts together with its gripper.



(b) Robot 2 pushes contacts into position.

Figure 15: Generation functions for the two robots.

5 Conclusions

We conclude with a list of the major highlights of this paper.

- Probabilistic Generation Theory is a generalization of the results in [3]. It is a simple yet comprehensive method that encompasses reproduction while now being capable of dealing with the probabilistic selection of resources. The existence of a von Neumann rank threshold below which degeneracy always occurs is demonstrated.

- A communication system may be viewed as a generation system and vice-versa. If we examine a generation system with the established results of Information Theory, we find that the von Neumann threshold of Information is in fact the channel capacity of the probabilistic generation process. It is only when this maximum amount of information is specified that the von Neumann system rank is achieved, and, realistically, no more information is required to completely specify an offspring machine. Furthermore, the reproduction rate proceeds at or below the von Neumann threshold of complexity per generation, and the full transfer of complexity implies that the offspring has the maximum amount of information needed to achieve the von Neumann system rank.

In addition to the lunar robotic application described in Section 4, current research efforts are also focused on the development of a seed identification algorithm to specify the minimal size for the seed of a viable robotic colony.

Acknowledgements

The authors are indebted to David Burke, Semyon Meerkov, Patrick Owens, and A. Galip Ulsoy for the insights that helped crystallize some of the notions in this paper.

References

[1] B. Adams and H. Lipson, “A universal framework for self-replication,” in *ECAL 2003*, ser. Lecture Notes in Computer Science, W. Banzhaf, T. Christaller, P. Dittrich, J. T. Kim, and J. Ziegler, Eds., vol. 2801. Springer-Verlag, 2003, pp. 1–9.

[2] R. A. Freitas Jr. and R. C. Merkle, *Kinematic Self-Replicating Machines*. Landes Bioscience, 2004. [Online]. Available: http://www.molecularassembler.com/KSRM.htm

[3] P. Kabamba, “The von neumann threshold of self-reproducing systems: Theory and computation,” The University of Michigan, Tech. Rep. CGR-06-11, 2006.

[4] P. Owens and A. G. Ulsoy, “Self-replicating machines: Preventing degeneracy,” The University of Michigan, Tech. Rep. CGR-06-02, 2006.

[5] C. E. Shannon, “A mathematical theory of communication,” *Bell System Technical Journal*, vol. 27, pp. 379–423 and 623–656, July and October 1948.

[6] M. Sipper, “Fifty years of research on self-replication: An overview,” *Artifical Life*, vol. 4, no. 3, pp. 237–257, 1998.

[7] J. von Neumann, *Theory of Self-Reproducing Automata*, A. Burks, Ed. University of Illinois Press, 1966.

Appendix

Proposition 1

Proof. The proof is by induction.

To initialize the induction, note that if $x \in M_1^\varepsilon$ then $x \in M_0$ by Definition 2. Therefore $M_0 \supseteq M_1^\varepsilon$.

Moreover, if $x \in M_2^\varepsilon$ then there exists $y \in M_1^\varepsilon$ such that $P[y = G(x,r)] > \varepsilon$.

Since $M_0 \supseteq M_1^\varepsilon$, then $y \in M_0$. This implies that $x \in M_1^\varepsilon$, which must mean that $M_1^\varepsilon \supseteq M_2^\varepsilon$.

For the induction hypothesis, assume $M_{k-1}^\varepsilon \supseteq M_k^\varepsilon$.

To prove by induction, if $x \in M_{k+1}^\varepsilon$ then there exists $y \in M_k^\varepsilon$ such that $P[y = G(x,r)] > \varepsilon$.

Since $M_{k-1} \supseteq M_k^\varepsilon$, then $y \in M_{k-1}$. This implies that $x \in M_k^\varepsilon$, which must mean that $M_k^\varepsilon \supseteq M_{k+1}^\varepsilon$. □

Proposition 2

Proof. The proof is by contradiction.

Suppose $\exists\, y \in M_i^\varepsilon$ for which $P[y = G(x,r)] > \varepsilon$, $\varepsilon > 0$.

Then, by definition of M_{i+1}^ε, $x \in M_{i+1}^\varepsilon$. But this contradicts the hypothesis that $x \in M_i^\varepsilon \backslash M_{i+1}^\varepsilon$.

Therefore, if $x \in M_i^\varepsilon \backslash M_{i+1}^\varepsilon$ and $P[y = G(x,r)] > \varepsilon$, then $y \notin M_i^\varepsilon$. □

Proposition 3

Proof. The proof is by induction.

To initialize the induction, the hypothesis states that $M_i^\varepsilon = M_{i+1}^\varepsilon$. Thus for $j = i+1$, $M_j^\varepsilon = M_i^\varepsilon$.

For the induction hypothesis, assume that $M_i^\varepsilon = M_{i+1}^\varepsilon = \ldots = M_{j-1}^\varepsilon = M_j^\varepsilon$.

To prove by induction, let $x \in M_j^\varepsilon$. Then by the definition of M_j^ε, x is ε-capable of producing $y \in M_{j-1}^\varepsilon$.

By the induction hypothesis, $M_{j-1}^\varepsilon = M_j^\varepsilon$. This implies that y is also $\in M_j^\varepsilon$, and hence $x \in M_{j+1}^\varepsilon$. Thus, $M_{j+1}^\varepsilon \supseteq M_j^\varepsilon$.

But by Proposition 1, $M_j^\varepsilon \supseteq M_{j+1}^\varepsilon$. This must mean that $M_{j+1}^\varepsilon = M_j^\varepsilon = M_i^\varepsilon$.

Thus if $M_i^\varepsilon = M_{i+1}^\varepsilon$, then for all $j \geq i$, $M_j^\varepsilon = M_i^\varepsilon$. □

Proposition 4

Proof. Let $\rho^\varepsilon(\Gamma) = \rho$ be finite. Then M can be decomposed as the union of mutually disjoint sets:

$$M = (M_0 \backslash M_1^\varepsilon) \cup (M_1^\varepsilon \backslash M_2^\varepsilon) \cup \cdots \cup (M_{\rho-1}^\varepsilon \backslash M_\rho^\varepsilon) \cup M_\rho^\varepsilon \quad (15)$$

The sets $(M_0 \backslash M_1^\varepsilon)$, $(M_1^\varepsilon \backslash M_2^\varepsilon)$, ..., $(M_{\rho-1}^\varepsilon \backslash M_\rho^\varepsilon)$ all contain at least one element (by the nature of Definition 2), so their respective cardinalities are lower-bounded by one. The cardinality of M_ρ^ε is lower-bounded by zero.

Taking the cardinalities of both sides of (15) yields $\rho^\varepsilon(\Gamma) \leq |M|$ for finite $\rho^\varepsilon(\Gamma)$.

Suppose however, that the rank is infinite. Then the sets $(M_0 \backslash M_1^\varepsilon)$, $(M_1^\varepsilon \backslash M_2^\varepsilon)$, ..., $(M_{|M|}^\varepsilon \backslash M_{|M|+1}^\varepsilon)$ are mutually disjoint. Again, each of these sets has to contain at least one element by Definition 2, so each of their respective cardinalities is lower-bounded by one. Now

$$M \supseteq (M_0 \backslash M_1^\varepsilon) \cup (M_1^\varepsilon \backslash M_2^\varepsilon) \cup \cdots \cup (M_{|M|}^\varepsilon \backslash M_{|M|+1}^\varepsilon) \quad (16)$$

and taking the cardinalities of both sides again gives the contradiction $|M| \geq |M| + 1$.

Therefore the rank cannot be infinite if there are a finite number of machines. The rank has to be less than or equal to the cardinality of the machine set. □

Proposition 5

Proof. The proof is by contradiction. Assume that any one of $x_i \notin M_\infty^\varepsilon$. Without loss of generality, let $x_i = x_1$. By Proposition 2, $\rho^\varepsilon(x_2) < \rho^\varepsilon(x_1)$, $\rho^\varepsilon(x_3) < \rho^\varepsilon(x_2)$, ..., $\rho^\varepsilon(x_n) < \rho^\varepsilon(x_{n-1})$, and therefore $\rho^\varepsilon(x_1) < \rho^\varepsilon(x_n)$. Then $\rho^\varepsilon(x_1) < \rho^\varepsilon(x_1)$, which is a contradiction. Hence if x_1, x_2,..., x_n form an ε-generation cycle of order n, then $x_i \in M_\infty^\varepsilon$, where $1 \leq i \leq n$. □

Proposition 6

Proof. We know that $M_\rho^\varepsilon = M_{\rho+1}^\varepsilon = M_\infty^\varepsilon$.

If $x \in M_\infty^\varepsilon$ then $x \in M_{\rho+1}^\varepsilon$.

By the definition of $M_{\rho+1}^\varepsilon$, there exists $y \in M_\rho^\varepsilon$ such that $P[y = G(x,r)] > \varepsilon$.

Since $y \in M_\rho^\varepsilon$ then $y \in M_\infty^\varepsilon$. □

Proposition 7

Proof. The proof is by contradiction.

Since the hypothesis states that $y_1 \in M_{\rho-1}^\varepsilon \backslash M_\rho^\varepsilon$ can be ε-generated, there exists $x \in M_\rho^\varepsilon$ such that $P[y_1 = G(x,r)] > \varepsilon$.

Let $R_1 \triangleq \{r_1 : y_1 = G(x, r_1)\}$.

Then $R_1 \neq \oslash$ because $P[y_1 = G(x,r)] > \varepsilon > 0.5$.

For the same x, and since ρ is finite, Proposition 6 implies that there exists $y_2 \in M_\rho^\varepsilon$ such that y_2 can be ε-generated.

Let $R_2 \triangleq \{r_2 : y_2 = G(x, r_2)\}$.

Then $R_2 \neq \oslash$ because $P[y_2 = G(x,r)] > \varepsilon > 0.5$.

We have $R_1 \subseteq R$; $R_2 \subseteq R$; and $(R_1 \cup R_2) \subseteq R$.

Since $y_1 \in M_{\rho-1}^\varepsilon \backslash M_\rho^\varepsilon$ and $y_2 \in M_\rho^\varepsilon$, then $y_1 \neq y_2$.

Now if $P[y_1 = G(x, r_1)] = P[R_1] > \varepsilon$, then for $y_2 \neq y_1$, $P[y_2 = G(x, r_2)] = P[R_2]$ should be $\leq 1 - \varepsilon$.

If $R_1 = R_2$, $\varepsilon < P[R_1] \leq 1 - \varepsilon$.

But then this means that $0 \leq \varepsilon \leq 0.5$, which is a contradiction.

Therefore, R_1 and R_2 are distinct and non-empty. Thus the resource set, R, must contain at least two elements. □

Proposition 8

Proof. The proof is by induction.

To initialize the induction, if $x \in M_1^{\varepsilon_1}$, then there exists $y_1 \in M_0^{\varepsilon_1}$ such that $P[y_1 = G(x,r)] > \varepsilon_1$.

Alternatively, if $x \in M_1^{\varepsilon_2}$, then there exists $y_2 \in M_0^{\varepsilon_2}$ such that $P[y_2 = G(x,r)] > \varepsilon_2$.

But $M_0^{\varepsilon_1} = M_0^{\varepsilon_2} = M_0$.

Since $\varepsilon_1 < \varepsilon_2$, and substituting $M_0^{\varepsilon_1}$ for $M_0^{\varepsilon_2}$ in the above, we can amend the last if-statement to read:

If $x \in M_1^{\varepsilon_2}$, then there exists $y_2 \in M_0^{\varepsilon_1}$ such that $P[y_2 = G(x,r)] > \varepsilon_1$.

This means that $x \in M_1^{\varepsilon_1}$. Hence, $M_1^{\varepsilon_1} \supseteq M_1^{\varepsilon_2}$.

Now if $x \in M_2^{\varepsilon_1}$, then there exists $y_1 \in M_1^{\varepsilon_1}$ such that $P[y_1 = G(x,r)] > \varepsilon_1$.

Alternatively, if $x \in M_2^{\varepsilon_2}$, then there exists $y_2 \in M_1^{\varepsilon_2}$ such that $P[y_2 = G(x,r)] > \varepsilon_2$.

But $M_1^{\varepsilon_1} \supseteq M_1^{\varepsilon_2}$.

Since $\varepsilon_1 < \varepsilon_2$, and substituting $M_1^{\varepsilon_1}$ for $M_1^{\varepsilon_2}$ in the above, we can amend the last if-statement to read:

If $x \in M_2^{\varepsilon_2}$, then there exists $y_2 \in M_1^{\varepsilon_1}$ such that $P[y_2 = G(x,r)] > \varepsilon_1$.

This means that $x \in M_2^{\varepsilon_1}$. Hence, $M_2^{\varepsilon_1} \supseteq M_2^{\varepsilon_2}$.

For the induction hypothesis, assume $M_k^{\varepsilon_1} \supseteq M_k^{\varepsilon_2}$.

To prove by induction, if $x \in M_{k+1}^{\varepsilon_1}$, then there exists $y_1 \in M_k^{\varepsilon_1}$ such that $P[y_1 = G(x,r)] > \varepsilon_1$.

If $x \in M_{k+1}^{\varepsilon_2}$, then there exists $y_2 \in M_k^{\varepsilon_2}$ such that $P[y_2 = G(x,r)] > \varepsilon_2$.

Using the induction hypothesis, $M_k^{\varepsilon_1} \supseteq M_k^{\varepsilon_2}$.

Since $\varepsilon_1 < \varepsilon_2$, and substituting $M_k^{\varepsilon_1}$ for $M_k^{\varepsilon_2}$ in the above, we can amend the last if-statement to read:

If $x \in M_{k+1}^{\varepsilon_2}$, then there exists $y_2 \in M_k^{\varepsilon_1}$ such that $P[y_2 = G(x,r)] > \varepsilon_1$.

This means that $x \in M_{k+1}^{\varepsilon_1}$. Hence, $M_{k+1}^{\varepsilon_1} \supseteq M_{k+1}^{\varepsilon_2}$. □

Proposition 9

Proof. Consider the probabilistic generation system to be operating at some ε_1.

Let $\rho^{\varepsilon_1}(\Gamma) = \rho_{\varepsilon_1}$.

That is, a machine $x_1 \in M_{\rho_{\varepsilon_1}}^{\varepsilon_1}$ is either able to ε_1-replicate or be a part of an ε_1-generation cycle. The machine rank of x_1 at ε_1 is $\rho^{\varepsilon_1}(x_1) = \rho_{\varepsilon_1}$.

Now consider the probabilistic generation system to be operating at some ε_2, where $\varepsilon_1 < \varepsilon_2$.

Let $\rho^{\varepsilon_2}(\Gamma) = \rho_{\varepsilon_2}$.

That is, a machine $x_2 \in M_{\rho_{\varepsilon_2}}^{\varepsilon_2}$ is either able to ε_2-replicate or be a part of an ε_2-generation cycle.

By Proposition 8, $M_{\rho_{\varepsilon_2}}^{\varepsilon_1} \supseteq M_{\rho_{\varepsilon_2}}^{\varepsilon_2}$.

In other words, $x_2 \in M_{\rho_{\varepsilon_2}}^{\varepsilon_2} \Rightarrow x_2 \in M_{\rho_{\varepsilon_2}}^{\varepsilon_1}$, and x_2 is able to ε_1-replicate as well (or be a part of an ε_1-generation cycle).

Taking the machine rank of x_2 at ε_1, we get $\rho^{\varepsilon_1}(x_2) = \rho_{\varepsilon_2}$.

But for x_2 to be able to ε_1-replicate (or be a part of an ε_1-generation cycle), that must mean that $\rho_{\varepsilon_1} \leq \rho_{\varepsilon_2}$.

$\Rightarrow \rho^{\varepsilon_1}(\Gamma) \leq \rho^{\varepsilon_2}(\Gamma)$. □

Proposition 10

Proof. From Proposition 9, we know that $\varepsilon_1 < \varepsilon_2$ implies that $\rho^{\varepsilon_1}(\Gamma) \leq \rho^{\varepsilon_2}(\Gamma)$, that is, the M_i^{ε} sets either become bigger or stay the same as ε goes to 0. Also, with Free Generation, $\varepsilon = 0$ implies that $\rho^{\varepsilon}(\Gamma) = 0$, that is, $M_i^{\varepsilon} = M_0$ when $\varepsilon = 0$. Therefore, the following statement is true:

For all $\gamma > 0$, there exists $\delta(\gamma)$ such that $\|\varepsilon\| < \delta$ implies that $\|\rho^{\varepsilon}(\Gamma)\| < \gamma$.

This is equivalent to saying that as ε goes to 0, M_{∞}^{ε} also goes to M_0. □