STRING REPLACEMENT SYSTEMS

by
Harvey M. Friedman
Distinguished University Professor
of Mathematics, Philosophy, Computer Science Emeritus
Ohio State University
Columbus, Ohio
https://u.osu.edu/friedman.8/foundational-adventures/downloadable-manuscripts/
https://www.youtube.com/channel/UCdRdeExwKiWndBl4YOxBTEQ
September 18, 2022

ABSTRACT. We propose finite string replacement as a foundation for recursion theory and computational complexity. It can be viewed as a simplified kind of formal grammar (Post, Chomsky). This has the conceptual advantage of not relying on ad hoc choices in the specific models of computation. This purity also suggests possible theoretical interactions with chemistry and molecular biology.

1. Introduction
2. DTM Model
3. DSGM Model
4. String Replacement: {0,...,n}* by SRS(m)
5. String Replacement: {0,...,n}* by SRS(n+1)
6. Deterministic String Replacement: {0,...,n}* by SRS(n)
7. P,NP,PSPACE subsets of {0,...,n}*
8. Further Research

1. INTRODUCTION.

Recursion theory and computational complexity theory are normally founded using models of computation that involve several ad hoc choices. Of course, famously, it was shown early on that varying these ad hoc choices lead to the same notions. Some of this development requires some more elaborate constructions going beyond direct computational ideas. It also begs for a certain kind of completion that amounts to a "proof" of Church's Thesis - and this remains arguably elusive.

Here we wish to provide a foundation for recursion theory and computational complexity theory that avoids ad hoc choices and stays within the realm of transparent computation.

The basic vehicle that we use for this treatment is that of String Replacement Systems. A String Replacement System consists of a list of finitely many rules

$x_1 \to y_1$
...
$x_r \to y_r$

where $x_1,\ldots,x_r,y_1,\ldots,y_r$ are **nonempty** finite strings of nonnegative integers (N).

We write i for the nonnegative integer, but also sometimes for the string of length 1, (i). It is always clear from context which is intended.

For greater control, we define an SRS(n) to be a finite list of rules

$x_1 \to y_1$

...
$x_r \to y_r$

where $x_1,\ldots,x_r,y_1,\ldots,y_r \in \{0,\ldots,n\}^*$. Here n is a nonnegative integer and A* is the set of all **nonempty** finite strings from A.

Of particular importance are bit strings, or the elements of SRS(1). It isn't until section 6 (with n = 1) that we are able to use bit strings exclusively for our results.

Let α be a SRS(n). The execution sequences in α are nonempty finite or infinite sequences of strings $w_1,w_2,\ldots$ from $\{0,\ldots,n\}^*$, where each w_{i+1} is obtained from w_i by choosing a rule $x \to y$ in α and replacing some consecutive substring (which is a copy of) x in w_i by y. Thus there is nondeterminism in two ways. First in the choice of rule in α and second in the choice of position of the consecutive substring.

A terminal execution sequence is a finite execution sequence which cannot be properly extended to an execution sequence. I.e., no rule can be applied to the last term.

The acceptance set of α is the set of first strings of the various terminal execution sequences.

The SRS(n) acceptance sets are the acceptance sets for the various SRS(n).

In section 4 we show that the SRS(0) acceptance sets are Presburger. We also show that A* is a SRS(n) acceptance set provided $A \subseteq \{0,...,n\}$. Also that each SRS(n) acceptance set is an SRS(m) acceptance set, $n \leq m$.

In section 5, we show that the SRS(1) acceptance sets (the bit string acceptance sets) include a complete r.e. set. This is an immediate consequence of Theorem 5.9. Thus already bit string replacement is very powerful.

Ideally, we would like to have

1) The r.e. subsets of $\{0,...,n\}^*$ exactly the SRS(n) acceptance sets, $n \geq 1$

In section 4 we refute 1) no matter which $n \geq 0$ is chosen. Moreover, we show that every nonempty SRS(n) acceptance set has an element of length 1.

In this paper we prove modified forms of 1), not only with r.e., but with P, NP, PSPACE.

We start by proving in section 4 that

2) For all $n \geq 0$ there exists $m > n$ such that the r.e. subsets of $\{0,...,n\}^*$ are exactly the SRS(m) acceptance sets intersect $\{0,...,n\}^*$

In section 5 we prove

3) Let $n \geq 0$. $0 \in S \subseteq \{0,...,n\}^*$ is r.e. if and only if S is the acceptance set of some SRS(n+1) intersect $\{0,...,n\}^*$.

4) Let $n \geq 0$. $S \subseteq \{0,...,n\}^*$ is r.e. if and only if S is the intersection of two (of finitely many) SRS(n+1) acceptance sets.

which are both close approximations to 1) above.

In section 6 we introduce premier execution sequences. These are deterministic. Here at each stage we must apply some rule at the first possible position, and then choose the first such rule. The premier acceptance set of SRS(n) is the set of first terms of its terminal premier execution sequences. The premier SRS(n) acceptance sets are the terminal premier execution sequences of the SRS(n).

In section 6, we prove

5) For all n ≥ 1, the r.e. subsets of {0,...,n}* are exactly the positive Boolean combinations of the premier SRS(n) acceptance sets.

In section 7 we adapt section 5 to NP and PSPACE instead of just r.e. For NP subsets of {0,...,n}*, we use a polynomial bound on the length of terminal execution sequences of SRS(n+1). For PSPACE subsets of {0,...,n}*, we use a polynomial bound on the length of the strings in terminal execution sequences of SRS(n+1).

In section 7, we also adapt section 6 to P instead of just r.e. For P subsets of {0,...,n}*, n ≥ 1, we use a polynomial bound on the length of premier terminal execution sequences of SRS(n).

Of particular note is the case of Theorems 5.9, 5.11 with n = 0, and Theorems 6.15, 7.3 with n = 1. These results specifically address bit string replacement.

2. DTM MODEL

We start with a standard deterministic DTM model well known to generate the r.e. subsets of the finite strings in any finite alphabet via acceptance. This is precisely the model used in [HU79], section 7.8, Two-way Infinite Tape, and with the tiny difference that here we do not allow the empty string for initialization.

We use distinct tape symbols 0,...,n,...,m,B where B is the blank and 0,...,n are the input symbols. We use distinct states q_0,...,,q_r, where q_0 is the initial state and q_r is the final state. Here 0 ≤ n ≤ m and r ≥ 1.

A symbol tape is a function f:Z → {B,0,...,n} which is B at all but finitely many arguments. We think of f as reading f(i) at "tape square" i ∈ Z.

Only certain symbol tapes are used for initializing computation. Let x ∈ {0,...,n}*. tape[x] is the symbol tape that reads x at arguments 1,...,lth(x), and is B at all other arguments.

Thus the integers name the squares of a two way infinite tape, and f(i) is the tape symbol written on square i.

A global state consists of (f,q,i), where f is a symbol tape, q is a state, and i ∈ Z. The idea is that the reading head is reading tape square i on the symbol tape f, and the system is in state q.

The global state initialized by x ∈ {0,...,n,B}* is (tape[x],q_0,1). Thus on initialization, the reading head is reading the first symbol of x, and the state is q_0.

Computation proceeds from one global state to the next. Initialization by x ∈ {0,...,n}* requires that the initial global state is (tape[x],q_0,1). Thus we are viewing q_0 as the initial state.

The transition function H tells us how computation goes from one global state to the next.

$H:\{q_0,...,q_{r-1}\} \times \{B,0,...,m\} \to \{q_0,...,q_r\} \times \{B,0,...,m\} \times \{L,R\}$.

Note that q_r is not allowed as an argument. This is because q_r is the final state.

Suppose computation is in the global state (f,q,i), and H(q,f(i)) = (q',s,A), s ∈ {0,...,m,B}, A ∈ {L,R}. We transition to global state (f',q',i') provided

f' is the same as f except that at i it is s.

i' = i+1 if A = R; i-1 if A = L.

Computation terminates precisely when a global state is reached whose state q is the final state q_r.

The acceptance set is the set of all x ∈ {0,...,n}* such that computation initialized by x eventually terminates.

THEOREM 2.1. (standard) The acceptance sets for the various DTM with input symbols 0,...,n are exactly the r.e. subsets of {0,...,n}*.

3. DSGM MODEL

We now remove the use of any tape and work only with finite strings with the special barrier symbols < and >, and a pointer (corresponding to the reading head). We write DSGM = deterministic string model.

We use distinct tape symbols 0,...,n,...,m,B,<,>, where B is the blank and 0,...,n are the input symbols. We use distinct states $q_0,...,q_r$, where q_0 is the initial state and q_r is the final state. Here $0 \leq n \leq m$ and $r \geq 1$. We still call these tape symbols even though there is no tape. Note that this is the same as for DTM except that we have added two special tape symbols < and >.

A global state consists of (<x>,q,i), where $x \in \{0,...,m,B\}^*$, q is a state, and $1 \leq i \leq lth(x)$. We think of i as a pointer to the i-th symbol in x. Note that we never point to < or >.

Computation proceeds from one global state to the next. Initialization by $x \in \{0,...,n\}^*$ requires that the initial global state is (<x>,q_0,1).

The transition function H tells us how computation goes from one global state to the next.

$H:\{q_0,...,q_{r-1}\} \times \{0,...,m,B\} \rightarrow \{q_0,...,q_r\} \times \{0,...,m,B\} \times \{L,R\}$.

Suppose computation is in the global state (<x>,q,i). Suppose $H(q,x_i) = (q',s,A)$, $s \in \{0,...,m,B\}$, $A \in \{L,R\}$. We transition to global state (<x'>,q',i') according to the following.

We reset the i-th symbol of x to be s, referring to this as x'.

i. A = L and i > 1. Set i' = i-1.
ii. A = R and i < lth(x). Set i' = i+1.
iii. A = L and i = 1. Put B in front of x'. Set i' = 1.
iv. A = R and i = lth(x). Put B in back of x'. Set i' = lth(x)+1.

Note that H sometimes instructs us to read and write B, but H does not instruct us to read and write < or >. However, there is an implicit reading and writing of < and > through the quantity i - in particular using i = 1 and i = lth(x).

Computation terminates with a last global state precisely when the global state has the final state q_r.

The acceptance set of the DSGM is the set of all $x \in \{0,...,n\}^*$ such that we reach the final state q_r when initializing by x.

THEOREM 3.1. The acceptance sets for the various DSGM with input symbols 0,...,n are exactly the r.e. subsets of $\{0,...,n\}^*$.

Proof: Let DTM be given with tape symbols 0,...,n,...,m,B and states $q_0,...,q_r$. We create a DSGM, with the same acceptance set, with tape symbols 0,...,n,...,m,B,<,> and states $q_0,...,q_r$. Thus we have the input symbols and tape symbols, the same initial and final states, and the same states. Fix an input string $x \in \{0,...,n\}^*$. We want to arrange a tight relationship between the global states of DTM and DSGM.

During computation in DTM, we can track the interval of squares of tape that have been visited by the reading head. We read the symbols from left to right on this interval of squares of tape to form $y \in \{0,...,m,B\}^*$, with the reading head pointing to the i-th symbol of y, and where the state of the DTM is q. This results in a sequence of global states of DSGM, with the corresponding transition function H. QED

4. STRING REPLACEMENT: {0,...,n}* BY SRS(m)

We now introduce String Replacement Systems (SRS). Here we get rid of states. Thus we have no states, no tape, and no reading head or pointer. The most natural execution is nondeterministic, although in section 6 we use a deterministic execution.

Let $n \geq 0$. A SRS(n) is a finite list of rules

$x_1 \rightarrow y_1$
...
$x_r \rightarrow y_r$

where $x_1,...,x_r,y_1,...,y_r \in \{0,...,n\}^*$. Recall our convention to not allow the empty string in the A*.

An execution sequence in a SRS(n) is a finite or infinite sequence $w_0,w_1,...$ from $\{0,...,n\}^*$, where each w_{i+1} is obtained from w_i in the following nondeterministic way. Some rule $x \rightarrow y$ is chosen, where x is a consecutive substring of w_i. There may be more than one such rule and also x may be a consecutive substring of w_i in more than one way. w_{i+1} is the result of replacing x by y.

A terminal execution sequence (tes) of a SRS(n) is an execution sequence where no rule applies to its last term.

The acceptance set of a SRS(n) is the set of all $x \in \{0,...,n\}^*$ which begins some tes.

The SRS(n) acceptance sets are the acceptance sets for the various SRS(n).

We begin by showing that the SRS(0) acceptance sets are Presburger. I.e., of the form $\{(0)^i: i \in A\}$ where $A \subseteq Z^+$ is Presburger.

THEOREM 4.1. Every SRS(0) acceptance set is Presburger.

Proof: Fix an SRS(0), and let S be its acceptance set. We argue by cases, where the first applicable case is operative.

case 1. There are no rules. $S = \{0\}^*$.

case 2. There is a rule $\{0\}^i \to \{0\}^i$. Then $S = \emptyset$.

case 3. There is a rule $\{0\}^i \to \{0\}^j$ where $i > j$, but no rule $\{0\}^i \to \{0\}^j$ where $i < j$. Then $S = \{0\}^*$.

case 4. There is a rule $\{0\}^i \to \{0\}^j$ where $i < j$, but no rule $\{0\}^i \to \{0\}^j$ where $i > j$. Then $S = \emptyset$.

case 5. There are rules $\{0\}^i \to \{0\}^j$, $\{0\}^{i'} \to \{0\}^{j'}$, where $i < j$ and $i' > j'$. Let i,i' be least possible. Let $r = (j-i)(i'-j')$.

We now claim that for all $n \geq i,i'$, $n \in S \leftrightarrow n+r \in S$. Let $n \geq i,i'$.

Suppose $n \in S$. Let α be a tes starting with n. We can start with n+r and successively go down to n via rule $\{0\}^{i'} \to \{0\}^{j'}$, and then copy α, to obtain a tes starting with n+r. Thus $n+r \in S$.

Suppose $n+r \in S$. Let β be a tes starting with n+r. We can start with n and successively go up to n+r via rule $\{0\}^i \to \{0\}^j$, and then copy β, to obtain a tes starting with n. Thus $n \in S$.

It is clear that these five cases are inclusive and that in each case, S is Presburger. QED

We shall see in section 5 that the SRS(1) acceptance sets include complete r.e. subsets of {0,1}*. Thus bit string replacement is very powerful.

THEOREM 4.2. For all A ⊆ {0,...,n}, A* is an SRS(n) acceptance set. For all 0 ≤ n ≤ m, every SRS(n) acceptance set is an SRS(m) acceptance set. Every SRS(n) acceptance set is an r.e. subset of {0,...,n}*.

Proof: Let A,n be as given. use the rules i → i for i ∈ {0,...,n}\A. Suppose x ∈ A*. Then x is terminal, and therefore accepted. Suppose x ∈ {0,...,n}*\A*. Let i ∈ {0,...,n}\A appear in x. Then x is not terminal and x transitions only to x. Therefore x is not accepted.

Let 0 ≤ n ≤ m and SRS(n) be given. Add the rules i → i for n < i ≤ m. Then this SRS(m) has the same acceptance set as SRS(n).

The third claim is clear since acceptance sets are existentially defined in the appropriate sense. QED

THEOREM 4.3. Every nonempty SRS(n) acceptance set, n ≥ 0, contains a string of length 1.

Proof: Suppose none of the length 1 strings 0,...,n are accepted. Then each of 0,...,n is not terminal, and hence every x ∈ {0,...,n}* is nonterminal, because we can executive some rule on the first letter. Hence the SRS(n) acceptance set is empty. QED

It will be convenient to use SRS(A) for SRS with an arbitrary finite alphabet A in place of just alphabets {0,...,n}.

THEOREM 4.4. For all n ≥ 0 there exists m > n such that the r.e. subsets of {0,...,n}* are exactly the SRS(m) acceptance sets intersect {0,...,n}*.

Proof: Let S ⊆ {0,...n}* be r.e. By Theorem 3.1, fix a DSGM with tape symbols 0,...,n,...,m,B,<,> and states $q_0,...,q_r$, whose acceptance set is S. As usual, the input symbols of the DSGM are 0,...,n, and the initial and final states of the DSGM are q_0 and q_r.

Let $A = \{0,\ldots,m,B,0',\ldots,n',<,>,],\$,q,q_0,\ldots,q_r\}$. We present our SRS(A). Note that we have added auxiliary symbols $0',\ldots,n',],\$,q,q_0,\ldots,q_r$.

We use the following rules for SRS(A).

GROUP 1

$0 \to <q0]$
...
$n \to <qn]$

GROUP 2

$]0 \to 0]$
...
$]n \to n]$

GROUP 3

$0\$ \to \$0'$
...
$n\$ \to \$n'$
$q\$ \to q_0$
$] \to \$>$

GROUP 4

$s< \to s<, s \in A$
$>s \to >s, s \in A$

Let the transition function of DSGM be $H:\{q_0,\ldots,q_{r-1}\} \times \{0,\ldots,m,B,<,>\} \to \{q_0,\ldots,q_r\} \times \{0,\ldots,m,B,<,>\} \times \{L,R\}$.

GROUP 5

For $H(q_i,s) = (q_j,t,L)$:

$vq_is \to q_jvt$
$v \in \{0,\ldots,n,B\}$
$<q_is \to <q_iBt$

where we replace 0,...,n by 0',...,n'

GROUP 6

For $H(q_i,s) = (q_j,t,R)$:

$q_isv \rightarrow tq_jv$
$v \in \{0,...,n,B\}$
$q_is> \rightarrow tq_jB>$
where we replace 0,...,n by 0',...,n'

Let $x \in \{0,...,n\}^*$, lth(x) = k. We show that x is accepted by the above SRS(A) if and only if x is accepted by the given DSGM. Let $x' \in \{0',...,n'\}^*$ be the result of replacing each i by i'.

The idea is for Group 1-4 execution to take us from the initial x to <q_0,x'>, in preparation for simulation of the DSGM via groups 5,6. Here the DSGM states are planted in the strings at the location just before the location of the DSGM pointer, with <,> used to mark front and back. Recall that DSGM pointers never point to < or >, and what is between <,> has no <,>.

CLAIM 1. In any execution sequence, if < appears after the first position of an entry, or > appears before the last position of an entry, then < will henceforth always appear after the first position in all later entries, or > will henceforth always appear before the last position in all later entries, respectively.

This is clear from inspecting the rules. The rules in group 4 don't do anything. Rules whose antecedent doesn't mention < (>) preserve these two properties. This leaves only the second rules within groups 5,6. These obviously also preserve these two properties.

CLAIM 2. In any terminal execution sequence in SRS(A) starting with x, the first rule applied is from group 1 at the first position, resulting in < at the first position. Hence by claim 1, all entries begin with <. Thereafter no group 1 rule is ever applied again. We arrive eventually at $\langle qx\$\rangle$. Henceforth, all entries begin with < and end with >, and we arrive at <q_0x'>. Thereafter only group 5,6 rules are applied. Henceforth group 5,6 rules get deterministically applied and we obtain the standard simulation of the given DSGM reaching termination with state q_r. x is accepted by the DSGM.

Let x,w_1,...,w_p be a terminal execution sequence in SRS(A). Clearly the only rules applicable to x are in group 1 as antecedents of other rules are not restricted to 0,...,n. If we apply at position > 1 to x then we get < after the first position, and according to claim 1, < appears after the first position in w_p. This violates that the sequence is terminal because of the applicability of group 4 to w_p. So group 1 is first applied at the first position, with w_1 = <qx_1]x_2...x_k. By inspection, < must begin each of w_1,...,w_p. It then follows that we cannot be applying group 1 rules after the first rule application.

Only a group 2 rule or the last group 3 rule can apply to w_1. We claim that we never use] → $> unless] is at the end. For if we use] → $> prematurely, > will appear before the end. By claim 1, > would appear before the end in w_p, and this would violate termination of the tes by the applicability of group 4 to w_p.

Hence we must be applying a group 2 rule until] is at the end, thus arriving at <qx$>. Now only group 3 rules can apply having ruled out group 1 above. These must be applied successively until we arrive at <q$x'>. Then we must apply the second to last rule in group 3 arriving at <q_0x'>. Since only group 5,6 rules have antecedent containing any q_i, $0 \le i < r$, we can now only be applying group 5,6 rules, and inspection shows that these applications are entirely deterministic. Furthermore, it is clear that the starting global state of the given DSGM with input string x is (<x>,q_0,1), and the corresponding entry we have arrived at is <q_0x'>. Clearly execution proceeds deterministically exactly as in DSGM with global states (<y>,q_i,j) in tandem with entries <y_1,...,$y_{j-1}q_iy_j$...y_k>, where 0,...,n is replaced by 0',...,n'. Thus since w_p is terminal, it must correspond to a global state with state q_r. Hence x is accepted in the DSGM.

Conversely, suppose x ∊ {0,...,n}* is accepted by DSGM. Then we proceed from x to <q_0x'> as above. Then as before, apply the group 5,6 rules arriving with q_r, where no rule can apply. In particular, group 1-3 rules don't apply because we have converted the 0,...,n to 0',...,n'. Also group 4 don't apply because < and > are at the first and last position.

Also note that x ∊ {0,...,n}* is accepted by DSGM if and only if x is accepted by this SRS(A) if and only if there is an execution sequence from x to a string containing q_r. QED

We need a modified construction where we avoid using (0) for the antecedent of any rule.

DEFINITION 4.1. SRS(m) is regular if and only if (0) is not the antecedent or the consequent of any of its rules. A regular SRS(m) acceptance set is the acceptance set of some regular SRS(m).

THEOREM 4.5. For all n ≥ 0 there exists m > n such that every r.e. subset of {0,...,n}* containing (0) is a regular SRS(m) acceptance set intersect {0,...,n}*.

Proof: Let S ⊆ {0,...n}* be r.e. containing (0). By Theorem 3.1, fix a DSGM with tape symbols 0,...,n,...,m,B,<,> and states q_0,...,q_r, whose acceptance set is S. As usual, the input symbols of the DSGM are 0,...,n, and the initial and final states of the DSGM are q_0 and q_r.

Let A = {0,...,m,B,0',...,n',<,>,],$,q,$q_0$,...,$q_r$}. We present our modified SRS(A). Note that we have added auxiliary symbols 0',...,n',],$,q,$q_0$,...,$q_r$.

We use the following rules for SRS(A)'.

GROUP 1a

1 → <q1]
... s
n → <qn]

GROUP 1b

0i → <q0i]
0 ≤ i ≤ n

GROUPS 2-6
same

]0 → 0]
...
]n → n]

Let x ∈ {0,...,n}*, where x does not start with 0. The proof of Theorem 4.4 goes thru for x without change. Now suppose x ∈ {0,...,n}* starts with 0 and is not (0). Then the proof of

Theorem 4.4 also applies, except that instead of having to start with 0 $\rightarrow$ <q0] from SRS(A), we start with the 0i $\rightarrow$ <q0i] from group 1b in SRS[A]', where x starts with 0i. Also note that (0) is accepted by SRS(A)'. QED

5. STRING REPLACEMENT: {0,...,n}* BY SRS(n+1)

We show that for every r.e. set A $\subseteq$ {0,...,m}*, n < m, containing (0) there exists SRS(n+1) accepting the same elements of {0,...,n}*.

According to Theorem 4.5, we fix regular SRS(m) whose acceptance set intersect {0,...,n}* is A.

It will be convenient to use \$ for n+1, so that SRS(n+1) = SRS(0,...,n,\$) = SRS(n,\$). We prefer the notation SRS(n,\$).

The idea is to interpret the letters 0,...,m, by certain strings α(1),...,α(m) $\in$ {0,,...,n,\$}*. The letters 0,...,n are left untouched. SRS(m) gets transformed into SRS(n,\$). We show α preserves acceptance of strings.

DEFINITION 5.1. Let x be a finite sequence from anywhere. We write lth(x) for the length of x, and x_i for the i-th term of x. x_i is defined if and only if $1 \leq i \leq$ lth(x). A block in x is a nonempty sequence of consecutive terms of x. An initial block in x is a block in x that starts with x_1. A final block in x is a block in x start ends with x[lth(x)].

DEFINITION 5.2. For $0 \leq i \leq n$, $\alpha(i) = i$. For $n+1 \leq i \leq m$, $\alpha(i)$ = \$\$0\$...0\$\$\$, where there are i many 0\$. For x $\in$ {0,...,m}*. $\alpha(x) = \alpha(x_1)\ldots\alpha(x_k)$, where lth($\alpha$) = k $\geq$ 1. Thus α:{0,...,m}* $\rightarrow$ {0,...,n,\$}*.

DEFINITION 5.3. Let R be the rule x $\rightarrow$ y in SRS(m). α(R) is the rule α(x) $\rightarrow$ α(y). SRS(n,\$) consists of the rules α(R) for rules R in SRS(m).

LEMMA 5.1. α:{0,...,m}* $\rightarrow$ {n,\$}* is one-one and the identity on {0,...,n}.

Proof: We prove the following by induction t. For all $i_1,\ldots,i_r,j_1,\ldots,j_s \in$ {0,...,m}, r,s $\leq$ t, if $\alpha(i_1,\ldots,i_r) = \alpha(j_1,\ldots,j_s)$ then $(i_1,\ldots,i_r) = (j_1,\ldots,j_s)$, Assume this is true for t. Now let $i_1,...,i_r,j_1,...,j_s \in \{0,...,m\}$, $r,s \leq t+1$ and $\alpha(i_1,...,i_r) = \alpha(j_1,...,j_s)$.

case 1. $i_1 \in \{0,...,n\}$. Then $j_1 \in \{0,...,n\}$, and therefore $i_1 = j_1$. Hence $\alpha(i_2,...,i_r) = \alpha(j_2,...,j_s)$, and so $(i_2,...,i_r) = (j_2,...,j_s)$ by the induction hypothesis.

case 2. $i_1 \in \{n+1,...,m\}$. Then $j_1 \in \{n+1,...,m\}$. Hence $\alpha(i_1)$ is an initial segment of $\alpha(j_1)$ or vice versa. Then obviously $i_1 = j_1$, and we can again apply the induction hypothesis.

QED

LEMMA 5.2. Suppose B is a block in $\alpha(x)$. Then
i. B is a proper block in some $\alpha(x_i)$; or
ii. B is some $\alpha(x_i)...\alpha(x_j)$, $i \leq j$; or
iii. B is some $\beta\gamma$, where exists i such that β is a proper final block in $\alpha(x_i)$ and γ is a proper initial block in $\alpha(x_{i+1})$; or
iv. B is some $\beta\alpha(x_i)...\alpha(x_j)$, $i \leq j$, β a proper final block in $\alpha(x_{i-1})$; or
v. B is some $\alpha(x_i)...\alpha(x_j)\gamma$, $i \leq j$, γ a proper initial block in $\alpha(x_{j+1})$; or
vi. B is some $\beta\alpha(x_i)...\alpha(x_j)\gamma$, $i \leq j$, β is a proper final block in $\alpha(x_{i-1})$ and γ a proper initial block in $\alpha(x_{j+1})$.

Proof: Suppose B is a block in $\alpha(x_1),...,\alpha(x_k)$. Look at how B is positioned. Let i be least such that B meets x_i and j be greatest such that B meets x_j. If I = j then we have i or ii. Assume i < j. There are four possibilities according to whether B contains x_i, B contains x_j. These correspond to iii - vi. QED

LEMMA 5.3. If x_k is a proper block in x_i then $k = 0$ and $n < i \leq m$.

LEMMA 5.4. Let β be a proper final block in α_i. For all k, α_k is not an initial block in β and β is not an initial block in α_k.

LEMMA 5.5. Let γ be a proper initial block in α_i. For all k, α_k is not a final block in γ and γ is not a final block in α_k.

LEMMA 5.6. Suppose $\alpha(y)$ is a block in $\alpha(x)$. Then $y = 0$ or y is a block in x.

Proof: Apply Lemma 5.2 to B = α(y) and α(x). If i then by Lemma 5.3, y = 0. By Lemma 5.4, iii,iv,vi is impossible. By Lemma 5.5, v is impossible. Hence y = 0 or ii holds. For ii, apply Lemma 5.1. QED

LEMMA 5.7. Let x,y $\in$ {0,...,m}*, and R be a rule in SRS(m). R applied to x is y if and only if α(R) applied to α(x) is α(y).

Proof: Let R read z $\rightarrow$ w. Since SRS(m) is regular, z,w $\neq$ (0). Suppose R applied to x is y. Then y is the result of replacing the block z in x by w. Hence α(y) is the result of replacing the block α(z) in α(x) by α(w).

Now suppose α(R) applied to α(x) is α(y). Let α(z) be the block in α(x) that is replaced by α(w) to form α(y), where R is z $\rightarrow$ w. By Lemma 5.6, z is a block in x, since z $\neq$ (0). We need to show that when we apply z $\rightarrow$ w to x we get y. Since z,w $\neq$ (0), we see that z is a block in w by Lemma 5.6. Hence α(R) applied to x is y. QED

LEMMA 5.8. Let x $\in$ {0,...,m}*.
i. x is terminal in SRS(m) if and only if α(x) is terminal in SRS(n,\$).
ii. x,y_1,...,y_r is a terminal execution sequence in SRS(m) if and only if α(x),α(y_1),...,α(y_r) is a terminal execution sequence in SRS(n,\$).
iii. For x $\in$ {0,...,m}*, x is accepted by SRS(m) if and only if α(x) is accepted by SRS(n,\$).
iv. (0) is terminal in SRS(n,\$).
v. For x $\in$ {0,...,n}*, x is accepted by SRS(m) if and only if x is accepted by SRS(n,\$).

Proof: Immediate from Lemma 5.7. QED

THEOREM 5.9. Let n $\geq$ 0. 0 $\in$ S $\subseteq$ {0,...,n}* is r.e. if and only if S is the acceptance set of some SRS(n+1) intersect {0,...,n}*.

Proof: By Lemma 5.8v. QED

LEMMA 5.10. Let n $\geq$ 0. {0,...,n}*\{0} is the acceptance set of some SRS(n+1).

Proof: We use the five rules

$n+1 \to n+1$
$0 \to 0$
$0i \to 1$
$i0 \to 1$
$0 \leq i \leq n$

Let n+1 appear in $x \in \{0,...,n+1\}^*$. All strings in all execution sequences from x have n+1, and hence none of them are terminal. So x is rejected.

0 can only go to 0 and hence 0 is rejected.

Suppose $x \in \{0,...,n\}^*$ is not 0. We prove by induction on j that if lth(x) = j then x is accepted. If j = 1 then x is terminal and so is accepted (if n = 0 then $j \neq 1$). Let x have length j+1. If there is no 0 in x then x is terminal, and hence accepted. Otherwise, let 0i appear in x or i0 appear in x. Then let x' be obtained by respectively replacing 0i and i0 by 1 (the third and fourth rules). Then x' has length j and is not 0. Hence by induction hypothesis x' is accepted. Since x yields x', clearly x is accepted. QED

THEOREM 5.11. Let $n \geq 0$. $S \subseteq \{0,...,n\}^*$ is r.e. if and only if S is the intersection of two (of finitely many) SRS(n+1) acceptance sets.

Proof: By Theorem 5.9 and Lemma 5.10. QED

We can't improve Theorem 5.9 by replacing n+1 by n.

THEOREM 5.12. $\{0,1,...,n\}$ is not the acceptance set of any SRS(n).

Proof: Let SRS(n) be as given. Then there must be a terminal string, and this terminal string must be of length ≥ 2. It is obviously accepted. QED

Can we improve Theorem 5.9 by eliminating the assumption $\{0\} \in S$? We leave this question open. However, the following is straightforward.

THEOREM 5.13. Every r.e. subset of $\{0,...,n\}^*$ is the acceptance set of some SRS(n+2) intersect $\{0,...,n\}^*$.

Proof: In proving Theorem 5.9, we use SRS(n,%,$} instead of SRS(n,$). So instead of using 0,$ for the $\alpha(x)$, $x \in \{0,...,m\}^*$, we use %,$ and eliminate the need for the assumption of 0 being included. We define $\beta(i) = i$ if $0 \leq I \leq n$, and $\beta(i)$ = $$%$...%$$$, where there are i many %$, $n < i \leq m$. Define $\beta(R)$ as before, where R is a rule of SRS(m). Define SRS(n,%,$) as the set of all $\beta(R)$. Here we are using % for n+1 and $ for n+2. Then prove that for all $x \in \{0,...,m\}^*$, $\beta(x)$ is accepted by SRS(n,%,$) if and only if x is accepted by SRS(m). QED

6. DETERMINISTIC STRING REPLACEMENT: {0,...,n}* by SRS(n)

Not every r.e. subset of {0,...,n}* is a SRS(n) acceptance set. This is clear from Theorems 4.3 and 5.10.

Here we will show that every r.e. subset of {0,...,n}* is a positive Boolean combination of what we call premier SRS(n) acceptance sets.

In a premier execution sequence, $w_1, w_2, ...$, we don't simply require that w_{i+1} is obtained from w_i by application of one of the rules somewhere in w_i. Instead we look for the first (leftmost) position that at least some rule can be applied, and then we apply the first listed rule at that first (leftmost) position. Of course, this does require that the SRS(n) have its rules listed without repetition. So far only the set of rules has been used.

Premier execution is obviously deterministic. As usual, termination occurs when no rule can be applied to an entry.

The premier acceptance set of a SRS(n) is the set of all $x \in \{0,...,n\}^*$ that starts a premiere terminal execution sequence for SRS(n).

The premier SRS(n) acceptance sets are the premier acceptance sets of the SRS(n).

LEMMA 6.1. Let $n \geq 2$. $\{x \in \{0,...,n\}^*: x_1 = 0\}$ is a premier SRS(n) acceptance set.

Proof: We use the following SRS(n).

$0i \rightarrow 0$

i ∈ {0,...,n}

i → i
i ∈ {0,...,m}\{0}

If x begins with i ≠ 0 then i → i is the premier application at the first position, and so i → i gets applied indefinitely, and x is not in the premier SRS(m) acceptance set.

If x begins with 0 then 0i → 0 is applicable at the front, removing the second letter, thereby reducing the length by 1. Either this results in 0 or the result is still the premier application. This continues until we are left with 0 alone, which is terminal. Hence x is in the premier SRS(n) acceptance set. QED

LEMMA 6.2. Let n ≥ 2 and S ⊆ {0,...,n}* be r.e. S and some premier SRS(n) acceptance set have the same elements that begin with 0.

Proof: Let m,S be as given. By Theorem 3.1, fix a DSGM with tape symbols 0,...,n,...,m,B,<,> and states q0,...,qr, whose acceptance set is S. As usual, the input symbols of the DSGM are 0,...,n, and the initial and final states of the DSGM are q_0 and q_r. Let A = {0,...,m,B,<,>,q_0,...,q_r} = {0,...,m+r+4}. This set equation indicates that, for example, we use > interchangeably with m+3.

We let f(i) = 12...21 and g(i) = 2...21, where 0 ≤ i ≤ m+r+4, and there are i+1 2's. We will be using the identification of A with {0,...,m+r+4}. f,g are extended to {0,...,m+r+4}* by concatenation.

The SRS(n) that we use is as follows.

GROUP 1

0 → f(0)11
f(i)11j → f(ij)11
0 ≤ i,j ≤ n

GROUP 2

g(i)11 → g(>i>)
0 ≤ i ≤ n

GROUP 3

$f(i>j) \rightarrow f(>ij)$
$i,j \leq n$

GROUP 4

$f(>j) \rightarrow f(<q_0j)$
$0 \leq j \leq n$

Let the transition function of DSGM be $H:\{q_0,\ldots,q_{r-1}\} \times \{0,\ldots,m,B\} \rightarrow \{q_0,\ldots,q_r\} \times \{0,\ldots,m,B\} \times \{L,R\}$.

GROUP 5

For $H(q_i,s) = (q_j,t,L)$:

$f(vq_is) \rightarrow f(q_jvs')$
$v \in \{0,\ldots,m,B\}$
$f(<q_is) \rightarrow f(<q_iBt)$

GROUP 6

For $H(q_i,s) = (q_j,t,R)$:

$f(q_isv) \rightarrow f(tq_jv)$
$v \in \{0,\ldots,m,B\}$
$f(q_is>) \rightarrow f(tq_jB>)$

Now fix $x \in A^*$ of length t for initialization, where $x_1 = 0$. We determine the premier $SRS(A) = SRS(m+r+4)$ sequence $x,w_1,w_2,\ldots$.

It is clear that the premier rule application to x is the first rule at position 1. Therefore $w_1 = f(0)11x_2\ldots x_t$. The first rule of Group 1 and the Group 2 rules can only be applied at a 0,2 respectively, and hence not at the first position of w_1, which begins with 1. The rule $f(0)11x_2 \rightarrow f(0x_2)11$ can apply to w_1 at the first position, and is the only such rule. Therefore $w_2 = f(0x_2)11x_3\ldots x_t$.

Since w_2 begins with 1 and has no $f(>)$, only groups 1,2 can apply to w_2. Since group 2 has antecedent with 111, it must apply at the second position of $f(x_2)$ in w_2. But group 1 line 2 applies at the first position of $f(x_2)$ in w_2. So the premier application must be from group 1 line 2, $f(x_2)11x_3 \rightarrow f(x_2x_3)11$, yielding $w_3 = f(0x_2x_3)11x_4...x_t$.

This must continue until we arrive at $w_t = f(0x_2x_3...x_t)11 = f(x)11$. This argument is valid even for t = 1.

Now $f(x)11 \in \{1,2\}^*$, and there is exactly one occurrence of 111, and it is at the very end with 2111. So group 1 line 2 cannot apply because that requires 2111j. Hence only group 2 applies, at the second position in the $f(x_n)$ in f(x)11, with $w_{t+1} = f(x_1...x_{t-1}>x_t>)$. Note that $w_{t+1} \in \{1,2,>\}^*$ and there are no 111 in w_{t+1}. Hence only groups 3,4 apply to w_{t+1}. The group 3 rules are applied at the beginning of the $f(x_{t-1})$, whereas the group 4 rules are applied at the beginning of the first f(>), which is later. Hence $w_{t+2} = f(x_1...>x_{t-1}x_t>)$. By the same reasoning, we continue with group 3, until we arrive at $w_{2t} = f(>x>)$.

Now Group 4 applies at the first position of w_{2t}, and we get $w_{2t+1} = f(<q_0x>)$. This argument is valid even if t = 1.

Since in w_{2t+1}, there are no 0's and no 111, and f(>) appears only at the end, only groups 5,6 can apply, and the result must also have no 0's, no 111, and f(>) appears only at the end. Hence from w_{2t+1} we are merely initializing with x and following DSGM with f applied. Therefore x is premier accepted by this SRS(n) if and only if x is accepted by DSGM(A) = DSGM(0,...,n+r+4). QED

LEMMA 6.3. Assume $n \geq 2$. Every r.e. subset of $\{x \in \{0,...,n\}^*: x_1 = 0\}$ is the intersection of two premier SRS(n) acceptance sets.

Proof: Let $S \subseteq \{x \in \{0,...,n\}^*: x_1 = 0\}$ be r.e. By Lemma 6.2, let A be a premier SRS(n) acceptance set that agrees with S on $\{x \in \{0,...,n\}^*: x_1 = 0\}$. Then $S = A \cap \{x \in \{0,...,n\}^*: x_1 = 0\}$. Now apply Lemma 6.1. QED

THEOREM 6.4. Assume $n \geq 2$. $S \subseteq \{0,...,n\}^*$ is r.e. if and only if S is a positive Boolean combination of premier SRS(n) acceptance sets.

Proof: Write $S = S_1 \cup ... \cup S_n$, where $S_i = \{x \in S: x_1 = i\}$. By Lemma 6.3 and symmetry, each S_i is the intersection of two premier $\{0,...,n\}^*$ acceptance sets. QED

We now want to handle the case n = 1.

LEMMA 6.5. Let S ⊆ {0,1}* be r.e. S and some premier SRS(1) acceptance set have the same elements that begin with 000.

Proof: This proof is almost the same as the proof of Lemma 6.2.

Let S be as given. By Theorem 2.1, fix a DSGM with tape symbols 0,...,n,...,m,B,<,> and states $q_0,...,q_r$, whose acceptance set is S. As usual, the input symbols of the DSGM are 0,...,n, and the initial and final states of the DSGM are q_0 and q_r, respectively. Let A = {0,...,m,B,<,>,q_0,...,q_r} = {0,...,n+r+4}. This set equation indicates that we use, for example, < interchangeably with m+2.

We let f(i) = 10...101 and g(i) = 0...101, where 0 ≤ i ≤ m+r+4, and there are i+1 copies of 10 in f(i), and f(i) = 1g(i). We will be using the identification of A with {0,...,m+r+4}. f,g are extended to {0,...,m+r+4}* by concatenation.

The SRS(1) that we use is as follows.

GROUP 1

000 → f(000)11
f(i)11j → f(ij)11
i,j ≤ 1

GROUP 2

g(i)11 → g(>i>)
i ≤ 1

GROUP 3

f(i>j) → f(>ij)
i,j ≤ 1

GROUP 4

f(>j) → f(<q_0j)
j ≤ 1

Let the transition function of DSGM be H:{q_0,...,q_{r-1}} × {0,...,m,B} → {q_0,...,q_r} × {0,...,m,B} × {L,R}.

GROUP 5

For $H(q_i,s) = (q_j,t,L)$:

$f(vq_is) \rightarrow f(q_jvs')$
$v \in \{0,...,m,B\}$
$f(<q_is) \rightarrow f(<q_iBt)$

GROUP 6

For $H(q_i,s) = (q_j,t,R)$:

$f(q_isv) \rightarrow f(tq_jv)$
$v \in \{0,...,m,B\}$
$f(q_is>) \rightarrow f(tq_jB>)$

Let $x \in \{0,1\}^*$ start with 000, for initialization, $lth(x) = t$. We argue that the premier execution of this SRS(1) terminates if and only if x is accepted by the DSGM(m). We argue as in the proof of Lemma 6.2. We determine the premier SRS(A) = SRS(m+r+4) sequence $x,w_1,w_2,...$.

It is clear that the premier rule application to x is the first rule at position 1. Therefore $w_1 = f(000)11x_2...x_t$. The first rule of Group 1 and the group 2 rules can only be applied at a 0, and hence not at the first position of w_1. The rule $f(0)11x_2 \rightarrow f(0x_2)11$ can apply to w_1 at the first position of the third f(0), whereas the group 2 rule can apply to w1 at the second position of the third f(0). Therefore $w2 = f(000x_2)11x_3...x_t$.

Since w_2 begins with 1 and has no f(>), only groups 1,2 can apply to w_2. Since group 2 has antecedent with 111, it must apply at the second position of $f(x_2)$ in w_2. But group 1 line 2 applies at the first position of $f(x_2)$ in w_2. So the premier application must be from group 1 line 2, $f(x_2)11x_3 \rightarrow f(x_2x_3)11$, yielding $w_3 = f(000x_2x_3)11x_4...x_t$.

This must continue until we arrive at $w_t = f(000x_2x_3...x_t)11 = f(x)11$. This argument is valid even for t = 1.

Note that in f(x)11 there is exactly one occurrence of 111, and it is at the very end with 0111. So group 1 line 2 cannot apply because that requires 0111j. Hence only group 2 applies, at the second position in the $f(x_n)$ in f(x)11, with $w_{t+1} = f(x_1...x_{t-1}>x_t>)$. Note that there are no 111 in w_{t+1}. Hence only groups 3,4 apply to wt+1. The group 3 rules are applied at the beginning of the $f(x_{t-1})$, whereas the Group 4 rules are applied at the beginning of the first f(>), which is later. Hence $w_{t+2} = f(x_1...>x_{t-1}x_t>)$. By the same reasoning, we continue with Group 3, until we arrive at $w_{2t} = f(>x>)$.

Now Group 4 applies at the first position of w_{2t}, and we get $w_{2t+1} = f(<q_0x>)$. This argument is valid even if t = 1.

Since in w_{2t+1}, there are no 111, and f(>) appears only at the end, only Groups 5,6 can apply, and the result must also have no 111, and f(>) appears only at the end. Hence from w_{2t+1} we are merely initializing with x and following DSGM with f applied. Therefore x is premier accepted by this SRS(1) if and only if x is accepted by DSGM(A) = DSGM(0,...,n+r+4). QED

LEMMA 6.6. Let S ⊆ {0,1}* be r.e. S and some premier SRS(2) acceptance set have the same elements that begin with 001.

Proof: We proceed as in the proof of Lemma 6.5 with the same f,g, and the slightly modified SRS(1). We only replace the first rule there by 001 → f(001)11. As in the proof of Lemma 6.5, we must start by applying 001 → f(001)11. Also no strings produced have 001, so that the first rule is never applied again. QED

LEMMA 6.7. Let S ⊆ {0,1}* be r.e. S and some premier SRS(1) acceptance set have the same elements that begin with 010.

Proof: We again proceed as in the proof of Lemma 6.6 but with a different f,g. We let f(i) = 10...01 and g(i) = 0...01, where there are i+1 0's. For the first rule, we use 010 → f(010)11. As in the proof of Lemma 6.5, we must start by applying 010 → f(010)11. Also no strings produced have 010, so that the first rule is never applied again. QED

LEMMA 6.8. Let i,j,k ≤ 1 and Let S ⊆ {0,1}* be r.e. S and some premier SRS(1) acceptance set have the same elements that begin with ijk.

Proof: By Lemmas 6.5 - 6.7, this is true for 000, 001, 010. Hence by symmetry this is true for 111, 110, 011, 101. QED

LEMMA 6.9. Let i,j,k ≤ 1. {x: x extends ijk} is a premier SRS(1) acceptance set.

Proof: Fix i,j,k ≤ 1. We use the following rules.

ijk → 0
i'j'k' → i'j'k'
i',j',k' ≤ 1
(i',j',k') ≠ (i,j,k)

First assume lth(x) ≥ 3. Clearly if x extends ijk then the premier transition from x is at the first position with the first rule. Hence x transitions to 0 which is obviously terminal. Therefore x is premier accepted. Suppose x does not extend ijk. Then the premier transition from x is at the first position with the unique rule after the first rule. This continues indefinitely, and therefore x is not premier accepted. QED

LEMMA 6.10. {(00),(0)} is a premier SRS(1) acceptance set.

Proof: We use the following rules.

ijk → ijk
i,j,k ≤ 1
1 → 1

Clearly 0 and 00 are premier accepted as they are terminal. No string of length ≥ 3 is premier accepted, via the first rule. Clearly 1,01,10,11 are not premier accepted, via the second rule. QED

LEMMA 6.11. {(00),(1)} is a premier SRS(1) acceptance set.

ijk → ijk
i,j,k ≤ 1
00 → 1
01 → 01
10 → 10
11 → 11
0 → 0

Two rules apply to 00 at first position, and the premier one yields 1. Clearly 1 is terminal. Hence 00 and 1 are premier accepted. Also strings of length ≥ 3 are not premier accepted, via the first rule. It remains to show that 01,10,11,0 are all not premier accepted. For 01, the premier transition is 01 → 01 and hence 01 is not premier accepted. Similarly for 10 and 11. Also 0 is obviously not premier accepted. QED

LEMMA 6.12. {(01),(0)} is a premier SRS(1) acceptance set.

Proof: We use the following rules.

ijk → ijk
i,j,k ≤ 1
01 → 0
00 → 00
1 → 1

One rule applies to 01 at first position, and this yields 0. Clearly 0 is terminal. Hence 0 and 01 are premier accepted. Also strings of length ≥ 3 are not premier accepted via the first rule. The only rule applying to 00 yields 00 and so 00 is not premier accepted. Also the rule applying to 10,11 yields 10,11, and so 10,11 are not premier accepted. 1 is obviously not premier accepted. QED

LEMMA 6.13. {(01),(1)} is a premier SRS(1) acceptance set.

Proof: We use the following rules.

ijk → ijk
i,j,k ≤ 1
01 → 1
10 → 10
00 → 00
11 → 11
0 → 0

The premier rule applying to 01 is 01 → 1 at first position, and this yields 1. Clearly 1 is terminal. Hence 1 and 01 are premier accepted. Also strings of length ≥ 3 are not premier accepted via the first rule. The rules applying to 00,10,11 yield 00,10,11, and so 00,10,11 are not premier accepted. 0 is obviously not premier accepted. QED

LEMMA 6.14. Every subset of $\{0,1\}^{\leq 2}$ is a positive Boolean combination of premier SRS(1) acceptance sets.

Proof: By taking the intersection of the sets in Lemmas 6.10, 6.11, we see this is true for {00}. By taking the intersection of the sets in Lemmas 6.12, 6.13, we see that this is true for {(01)}. By symmetry, this is true for {11}, {10}. Thus every subset of $\{0,1\}^{\leq 2}$ is a positive Boolean combination of premier SRS(1) acceptance sets.

We claim that {0} is the premier SRS(1) acceptance set for

$00 \rightarrow 00$
$1 \rightarrow 1$

For (0) is terminal and so premier accepted. Any string with 1 somewhere premier transitions to itself and so is not accepted. Any sting consisting of all zeroes, and two or more of them, premier transitions to itself, and so is not premier accepted.

By symmetry, {(1)} is a premier SRS(1) acceptance set. Thus we have shown that every singleton from $\{0,1\}^{\leq 2}$ is a positive Boolean combination of premier SRS(1) acceptance sets. Hence this is true of all subsets of $\{0,1\}^{\leq 2}$. QED

THEOREM 6.15. Let $n \geq 1$. $S \subseteq \{0,...,n\}^*$ is r.e. if and only if S is a positive Boolean combination of premier SRS(n) acceptance sets.

Proof: We have already proved this for $n \geq 2$ (Theorem 6.4), and so we set $n = 1$. Let $S \subseteq \{0,1\}^*$ be r.e. Let $u_1,...,u_8$ enumerate the bit strings of length 3 without repetition. Set $S_i = S \cap \{x: x \text{ extends } u_i\}$. By Lemma 6.8, Let S_i' be a premier SRS(1) acceptance set having the same elements of S_i that begin with u_i. Let $S_i'' = S_i' \cap \{x \in \{0,1\}^*: x \text{ extends } u_i\}$. Note that

$$S = S_1'' \cup ... \cup S_8'' \cup (S \cap \{0,1\}^{\leq 2})).$$

This writes S as a positive Boolean combination of premier SRS(n) acceptance sets using Lemmas 6.9, 6.14. QED

7. P,NP,PSPACE SUBSETS OF {0,...,n}*

Here we adapt sections 2-5 from r.e. subsets of {0,...,n}* to P,NP,PSPACE subsets of {0,...,n}*.

Sections 4,5 adapt to NP by putting a polynomial bound along with the SRS(m) or SRS(n+1), written SRS(m)/PL or SRS(n+1)/PL. Here acceptance is made stronger with the polynomial P, by requiring that the length of terminal execution sequences be bounded by P(lth(x)), where x is the initial string. We obtain

THEOREM 7.1. Let $n \geq 0$. $0 \in S \subseteq \{0,...,n\}^*$ is NP if and only if S is the intersection of an SRS(n+1)/PL acceptance set with $\{0,...,n\}^*$. Also $S \subseteq \{0,...,n\}^*$ is r.e. if and only if S is the intersection of two (of finitely many) SRS(n+1) acceptance sets.

The idea is to use the obvious nondeterministic forms of DTM and DSGM, and imitate the proofs of Theorems 5.11 and 5.13. In the proofs of Theorems 4.5, 5.11, 5.13, terminal execution always has a preparation phase of polynomial length, followed by simulation of the computation model, which in this case is NP, and therefore also of polynomial length.

Sections 4,5 adapt to PPACE by putting a polynomial bound along with the SRS(m) or SRS(n+1), written SRS/PS or SRS(n+1)/PS. Here acceptance is made stronger with the polynomial P, by requiring that the length of strings used in the execution sequences be bounded by P(lth(x)), where x is the initial term. We obtain

THEOREM 7.2. Let $n \geq 0$. $0 \in S \subseteq \{0,...,n\}^*$ is in PSPACE if and only if S is the intersection of an SRS(n+1)/PS acceptance set with $\{0,...,n\}^*$. Also $S \subseteq \{0,...,n\}^*$ is in PSPACE if and only if S is the intersection of two (of finitely many) SRS(n+1)/PS acceptance sets.

The idea is to keep the deterministic DTM and DSGM, and again imitate the proofs of Theorems 5.11 and 5.13. In the proofs of Theorema 4.5, 5.11, 5.13, terminal execution always has a preparation phase of polynomial length, followed by simulation of the computation model, which in this case is PSPACE, and therefore also with strings of polynomial length. These strings correspond to tape descriptions which are assumed to be of polynomial size.

Section 6 adapts to PTIME by putting a polynomial bound along with the SRS(n), written SRS(n)/PL. Here premier acceptance is made stronger with the polynomial P, by requiring that the length of terminal premier execution sequences be bounded by P(lth(x)), where x is the initial string. In the proof of Theorem 6.15, premier execution always has a preparation part of polynomial length, followed by a simulation of the computation model, which in this case is in P. The other Lemmas involve only direct polynomial length processes.

THEOREM 7.3. Let $n \geq 1$. $S \subseteq \{0,...,n\}^*$ is P if and only if S is a positive Boolean combination of premier SRS(n)/PL acceptance sets.

8. FURTHER RESEARCH

We don't have a good understanding of what the SRS(n) acceptance sets are, $n \geq 0$. We also don't have a good understanding of what the premier SRS(n) acceptance sets are, $n \geq 0$.

How can we improve the main results of sections 5-7?

How can we characterize the r.e., P, NP, and PSPACE subsets of {0,...,n}* in terms of the SRS(n) acceptance sets instead of going to SRS(n+1)?

How can we likewise treat the more refined complexity classes used in computational complexity theory, instead of the relatively crude r.e., P, NP, PSPACE?

The purity of the string replacement systems, especially bit string replacement systems, and perhaps even with premier execution, suggests possible theoretical interactions with chemistry and molecular biology. For instance, those sciences may suggest additional models of processing related to the present string replacement systems, worthy of analogous mathematical investigation.

REFERENCES

[Ch56] N. Chomsky, Three models for the description of language (PDF). IRE Transactions on Information Theory. **2** (3): 113-124.

[Ch59] N. Chomsky, On certain formal properties of grammars (PDF). Information and Control. **2** (2): 137-167.

[HU79] J. Hopcroft, J. Ullman, Introduction to Automata Theory, Languages, and Computation, 1979, Addison Wesley, 418 pages.

[Po36] E. Post, "Finite Combinatory Processes - Formulation 1". *Journal of Symbolic Logic*. 1 (3): 103–105.

[Po43] E. Post, "Formal Reductions of the General Combinatorial Decision Problem". *American Journal of Mathematics*. 65 (2): 197–215.