CDMTCS Research Report Series

An Explicit Construction of a Universal Extended H System

Gordon Alford
Department of Computer Science
University of Auckland
Auckland, New Zealand

CDMTCS-043
August 1997

Centre for Discrete Mathematics and
Theoretical Computer Science

An Explicit Construction of a Universal Extended H System

Gordon Alford

Department of Computer Science
University of Auckland
Private Bag 92019
New Zealand

Abstract

Lately there has been much interest concerning H systems, a generative mechanism based on the splicing operation, itself a language-theoretic equivalent of DNA recombination. Păun et al. have shown that regular extended H systems are theoretically universal but one has not yet been explicitly constructed. In this paper we explicitly construct a universal extended H system containing 182 axioms and 270 groups of rules.

1 Introduction

Molecular computing covers different models of computation involving operations on strands of DNA. As DNA is incredibly complex this potentially gives us a previously unobtainable degree of parallelization.
The study of H systems is a new branch of formal language theory and a significant theoretical component of molecular computing. H systems were first developed in 1987 by Tom Head [2] as a model of computation based upon the splicing operation, a language-theoretic model of DNA recombination. Extended H systems were then considered in 1996 by Păun et al. [5] and are the primary focus of this paper.
One important property within formal language theory is universality. Universality enables the comparison between various models of computability. It does this by considering the class of computable problems and determining whether or not a given model can generate a solution for any such problem. This is a fundamental characteristic for any computational model and especially relevant as regards H systems as evidenced by many of the recent results, in particular Păun [3] and Csuhaj-Varjú et al. [1].
Gheorghe Păun [4] posed us the following question :

Can we explicitly construct universal extended H systems of various types?

Theoretical results support this and it is the aim of this paper to offer such a construction where the resulting extended H system has a finite set of axioms and a regular set of rules.

2 Notation, Definitions and Previous Results

We denote by V^* the free monoid generated by the alphabet V, by λ the empty string and by V^+ the set $V^* - \{\lambda\}$

A *rewriting system* is a pair $\rho = (V, F)$ where V is an alphabet and F a finite set of ordered pairs of words over V.

A rewriting system $\tau = (V, F)$ is called a *Turing Machine* iff the following conditions are satisfied.

i) V is divided into two disjoint alphabets S and V_T, referred to as the *state* and *tape* alphabets.

ii) Elements $s_1 \in S$, $\# \in V_T$, and a subset $S_1 \subseteq S$ are specified, namely the *initial state*, the *boundary marker*, and the *final state set*. The set $V_1 = V_T - \{\#\}$ is not empty. An element $0 \in V_1$ and a subset $V_I \subseteq V_1$ are specified.

iii) The productions in F are of the forms

$$\begin{array}{rl} s_i a \to s_j b & \text{(overprint)} \\ s_i a c \to a s_j c & \text{(move right)} \\ s_i a\# \to a s_j 0\# & \text{(move right and extend workspace)} \\ c s_i a \to s_j c a & \text{(move left)} \\ \# s_i a \to \# s_j 0 a & \text{(move left and extend workspace)} \end{array}$$

where $s_i, s_j \in S$ and $a, b, c \in V_1$. Furthermore, for each $s_i, s_j \in S$ and $a \in V_1$, F either contains no productions of the second and third types or else contains both for every $c \in V_1$ (respectively for productions of the fourth and fifth types). Also for no $s_i \in S$ and $a \in V_1$ is the word $s_i a$ a subword of the left side of two productions of the first, third and fifth types.

We say that a word sP, where $s \in S$ and $P \in V_T^*$, is *final* iff P does not begin with a letter a such that sa is a subword of the left side of some production in F.
We define two Turing machines τ_1 and τ_2 to be *equivalent* iff $L(\tau_1) = L(\tau_2)$.
The language *accepted* by a Turing Machine τ is defined by

$$L(\tau) = \{P \in V_I^* \mid \#s_1P\# \Rightarrow^* \#P_1 s_i P_2\# \quad \text{for some } s_i \in S_1, \\ P_1, P_2 \in V_1^*, \text{ such that } s_i P_2\# \text{ is final}\}$$

A *analytic grammar* is a quadruple $G = (V_N, V_T, X_0, F_G)$ where V_N and V_T are disjoint alphabets, $X_0 \in V_N$, and F_G is a finite set of ordered pairs (u, v) such that u and v are words over the alphabet $V_N \cup V_T$ and v contains at least one letter of V_N. The elements of V_N are called *nonterminals* and those of V_T *terminals*. X_0 is called the *initial* letter and the elements of F_G are called *rewriting rules* or *productions* and are written as $u \to v$.
A grammar G with no restrictions, as given above, is called a type-0 grammar.
The language *accepted* by G is defined by

$$L(G) = \{P \mid P \in V_T^*,\ P \Rightarrow^* X_0\}$$

The following result is given in Salomaa [7]. The construction within the proof is used in the translation from a universal Turing machine to an equivalent type-0 grammar. Thus for completeness we include the proof of this result in our paper.

Theorem 2.1. *If a language is acceptable by a Turing machine τ, then it is of type-0.*

Proof. Assume that $L = L(\tau)$ where in connection with τ we use the notations of the definition. We define a type-0 analytic grammar G which recognizes L. The terminal alphabet of G is V_I. The nonterminal alphabet consists of the letter in $V - V_I$ and of the additional letters X_0, X_1 and X_2. The initial letter is X_0. The production set of G consists of the productions of τ and of the productions

$$\lambda \to \#s_1, \quad \lambda \to \#, \quad s_i a \to X_1, \quad X_1 b \to X_1,$$
$$X_1\# \to X_2, \quad s_i\# \to X_2, \quad bX_2 \to X_2, \quad \#X_2 \to X_0$$

where s_i ranges over S_1, b ranges over V_1, and for each s_i, a ranges over such elements of V_1 that $s_i a$ is final. It can now be verified that $L(G) = L(\tau)$.
If $P \in L(\tau)$, there is a derivation according to G where if $P = \lambda$

$$P \Rightarrow \#s_1 \Rightarrow \#s_1\# \Rightarrow \#X_2 \Rightarrow X_0$$

or alternatively if $P \neq \lambda$

$$P \Rightarrow \#s_1 P \Rightarrow \#s_1 P\# \Rightarrow^* \#P_1 s_i a P_2\# \Rightarrow \#P_1 X_1 P_2\# \Rightarrow^* \#P_1 X_2 \Rightarrow^* X_0$$

Consequently, $P \in L(G)$.
Assume, conversely, that $P \in L(G)$.
If $P = \lambda$, there is a derivation according to G from $\#s_1\#$ to X_0. Then $\lambda \in L(\tau)$.
If $P \neq \lambda$, there is a derivation according to G from $\#P_1 s_i a P_2\#$ to X_0, and a derivation from P to $\#P_1 s_i a P_2\#$ where $s_i \in S_1,\ a \in V_1,\ P_1, P_2 \in V_1^*$ such that $s_i a$ is final.
Thus $P \in L(\tau)$. □

An *extended H system* is a quadruple $\gamma = (V, T, A, R)$ where V is an alphabet, $T \subseteq V$, $A \subseteq V^*$, and $R \subseteq V^*\#V^*\$V^*\#V^*$, with $\#, \$$ special symbols not in V.
We call V the alphabet of γ, T the *terminal* alphabet, A the set of *axioms*, and R the set of *splicing rules*.
For $x, y, z \in V^*$ and $r : u_1\#u_2\$u_3\#u_4$ in R, we write

$(x, y) \vdash_r z$ iff $x = x_1u_1u_2x_2,\ y = y_1u_3u_4y_2$ and $z = x_1u_1u_4y_2$ for some $x_1, x_2, y_1, y_2 \in V^*$

With respect to an H system γ and a language $L \subseteq V^*$, we define

$$\sigma(L) = \{z \in V^* \mid (x, y) \vdash_r z \text{ for some } x, y \in L, r \in R\}$$

Then

$$\sigma^*(L) = \bigcup_{i \geq 0} \sigma^i(L) \quad \text{where} \quad \sigma^0(L) = L$$
$$\sigma^{i+1}(L) = \sigma^i(L) \cup \sigma(\sigma^i(L)), i \geq 0$$

The *language generated* by the H system γ is then defined by $L(\gamma) = \sigma^*(A) \cap T^*$

The following result appears in Păun [3]. The construction within the proof is used in the translation from a universal type-0 grammar to an equivalent universal extended H system. Thus for completeness we include an outline of the proof of this result in our paper.

Theorem 2.2. *The family of recursively enumerable languages coincides with the family of languages generated by extended H systems $\gamma = (V, T, A, R)$, where the set of axioms A is a finite language and the set of rules R is a regular language.*

Proof. Consider a type-0 grammar $G = (V_N, V_T, X_0, F_G)$ and construct the extended H system

$$\gamma = (V, T, A, R)$$

where

$V = V_N \cup V_T \cup \{X, X', B, Y, Z\} \cup \{Y_\alpha \mid \alpha \in V_N \cup V_T \cup \{B\}\}$
$T = V_T$
$A = \{XBX_0Y, ZY, XZ\} \cup \{ZvY \mid u \to v \in F_G\} \cup \{ZY_\alpha, X'\alpha Z \mid \alpha \in V_N \cup V_T \cup \{B\}\}$

and R contains the following groups of rules :

1) $Xw\#uY\$Z\#vY$ for $u \to v \in F_G,\ w \in (V_N \cup V_T \cup \{B\})^*$
2) $Xw\#\alpha Y\$Z\#Y\alpha$ for $\alpha \in V_N \cup V_T \cup \{B\},\ w \in (V_N \cup V_T \cup \{B\})^*$
3) $X'\alpha\#Z\$X\#wY_\alpha$ for $\alpha \in V_N \cup V_T \cup \{B\},\ w \in (V_N \cup V_T \cup \{B\})^*$
4) $X'w\#Y_\alpha\$Z\#Y$ for $\alpha \in V_N \cup V_T \cup \{B\},\ w \in (V_N \cup V_T \cup \{B\})^*$
5) $X\#Z\$X'\#wY$ for $w \in (V_N \cup V_T \cup \{B\})^*$
6) $\#ZY\$XB\#wY$ for $w \in T^*$
7) $\#Y\$XZ\#$

The rules in group 1 above encode only the productions of G. Groups 2-5 produce circular permutations of a string $Xw\alpha Y$ and nothing more, thus enabling the rules in group 1 to be applied at any place in a sentential form w of G. This allows any production of G to be simulated in γ. We now consider groups 6 and 7 but these will only produce terminating strings if they are applied sequentially, in order, in which case they will only give terminal forms of strings $XBwY$ where w is composed only of elements of T, hence $L(G) \subseteq L(\gamma)$, $L(\gamma) \subseteq L(G)$ and thus $L(G) = L(\gamma)$. □

As the symbol # is used as a marker for the rules of the H systems we shall denote by $T_\#$ the translation of the symbol # from either Turing machines or grammars to H systems.

3 Equivalent Turing Machines

As there are many ways of describing a given Turing machine we consider the equivalences between two descriptions and prove that they are equivalent.
The Turing machine that we consider is used in Rogozhin [6]. Productions are of the form q_ixyIq_j where $q_i, q_j \in S$, $x, y \in V_1$, $I \in \{L, M, R\}$ and can be read as: start in state q_i with symbol x, write symbol y, move in direction I and change into state q_j.

Let $\tau_r = (V_r, F_r)$ be a Turing machine of the type used in Rogozhin [6] and $\tau = (V, F)$ be a Turing machine as defined in section 2

Theorem 3.1. *Given an arbitrary Turing machine τ_r there exists an equivalent Turing machine τ.*

Proof. Consider a machine $\tau = (V, F)$. We then construct a machine τ_r :
Let $V_r = V$.
Now construct F_r from F :
If $P \in F$ is of the form $s_i a \to s_j b$ then define a new production $q_i abM q_j$ in F_r
If $P \in F$ is of the form $s_i ac \to as_j c$ then define a new production $q_i aaR q_j$ in F_r
If $P \in F$ is of the form $cs_i a \to s_j ca$ then define a new production $q_i aaL q_j$ in F_r
If $P \in F$ is of the form $s_i a\# \to as_j 0\#$ or $\# s_i a \to \# s_j 0a$ then no productions need to be added to F_r as there will be a $P' \in F$ of the form $s_i ac \to as_j c$ or $cs_i a \to s_j ca$ respectively.
Thus $L(\tau) \subseteq L(\tau_r)$
Consider a machine $\tau_r = (V_r, F_r)$. We then construct a machine τ :
Let $V_r = V_T \cup \{q_i \mid i \in 1..m\}$
Let $r_1..r_m$ be new states not in V_r.
Then $V = V_T \cup \{s_i \mid i \in 1..m\} \cup \{r_i \mid i \in 1..m\}$
Now construct F from F_r :
If $P \in F_r$ is of the form $q_i xyR q_j$ then define the following new productions in F :

$$\begin{array}{rl} s_i x \to r_i y & \text{(overprint)} \\ r_i yc \to ys_j c & \text{(move right)} \\ r_i y\# \to ys_j 0\# & \text{(move right and extend workspace)} \end{array}$$

If $P \in F_r$ is of the form $q_i xyM q_j$ then define a new production $s_i x \to s_j y$ in F
If $P \in F_r$ is of the form $q_i xyL q_j$ then define the following new productions in F :

$$\begin{array}{rl} s_i x \to r_i y & \text{(overprint)} \\ cr_i y \to s_j cy & \text{(move left)} \\ \# r_i y \to \# s_j 0y & \text{(move left and extend workspace)} \end{array}$$

And so τ fulfills the conditions of the definition.
Thus $L(\tau_r) \subseteq L(\tau)$ and so we have that $L(\tau) = L(\tau_r)$. □

The converse of the theorem also holds by the same argument.

4 An Explicit Universal H System

The universal Turing machine that we consider is $UTM(24, 2)$ described in Rogozhin [6].
Let $\tau_r = (V_r, F_r)$ be the $UTM(24, 2)$ where V_r and F_r are :
$V_r = \{0, 1, \#\} \cup \{q_i \mid i \in 1..24\}$

$$\begin{array}{llllllll} F_r = \{ & q_1 00Rq_5 & q_2 01Rq_1 & q_3 00Lq_4 & q_4 01Lq_{12} & q_5 01Rq_1 & q_6 00Lq_7 \\ & q_1 11Rq_2 & q_2 11Lq_3 & q_3 10Lq_2 & q_4 10Lq_9 & q_5 10Lq_6 & q_6 11Lq_7 \\ & q_7 00Lq_8 & q_8 00Lq_7 & q_9 00Rq_{19} & q_{10} 01Lq_4 & q_{11} 00Lq_4 & q_{12} 00Rq_{19} \\ & q_7 10Lq_6 & q_8 11Rq_2 & q_9 11Lq_4 & q_{10} 10Rq_{13} & q_{11} 1- & q_{12} 11Lq_{14} \\ & q_{13} 00Rq_{10} & q_{14} 00Lq_{15} & q_{15} 00Rq_{16} & q_{16} 00Rq_{15} & q_{17} 00Rq_{16} & q_{18} 00Rq_{19} \\ & q_{13} 11Rq_{24} & q_{14} 11Lq_{11} & q_{15} 11Rq_{17} & q_{16} 11Rq_{10} & q_{17} 11Rq_{21} & q_{18} 11Rq_{20} \\ & q_{19} 01Lq_3 & q_{20} 01Rq_{18} & q_{21} 00Rq_{22} & q_{22} 01Lq_{10} & q_{23} 01Rq_{21} & q_{24} 00Rq_{13} \\ & q_{19} 11Rq_{18} & q_{20} 10Rq_{18} & q_{21} 11Rq_{23} & q_{22} 11Rq_{21} & q_{23} 10Rq_{21} & q_{24} 10Lq_3 \; \} \end{array}$$

Using Theorem's 2.1 & 3.1 to transform $\tau_r = (V_r, F_r)$ into a type-0 grammar G gives :

$V_N = \{X_0, X_1, X_2\}$, $V_T = \{0, 1, \#\} \cup \{s_i \mid i \in 1..24\}$ and

$F_G = \{\lambda \to \#$	$X_1 0 \to X_1$	$X_1 1 \to X_1$	$X_1 \# \to X_2$
$\lambda \to \# s_1$	$0 X_2 \to X_2$	$1 X_2 \to X_2$	$\# X_2 \to X_0$
$s_{11} 1 \to X_1$			
$s_1 \# \to X_2$	$s_2 \# \to X_2$	$s_3 \# \to X_2$	$s_4 \# \to X_2$
$s_5 \# \to X_2$	$s_6 \# \to X_2$	$s_7 \# \to X_2$	$s_8 \# \to X_2$
$s_9 \# \to X_2$	$s_{10} \# \to X_2$	$s_{11} \# \to X_2$	$s_{12} \# \to X_2$
$s_{13} \# \to X_2$	$s_{14} \# \to X_2$	$s_{15} \# \to X_2$	$s_{16} \# \to X_2$
$s_{17} \# \to X_2$	$s_{18} \# \to X_2$	$s_{19} \# \to X_2$	$s_{20} \# \to X_2$
$s_{21} \# \to X_2$	$s_{22} \# \to X_2$	$s_{23} \# \to X_2$	$s_{24} \# \to X_2$
$s_1 00 \to 0 s_5 0$	$s_1 01 \to 0 s_5 1$	$s_1 0 \# \to 0 s_5 0 \#$	$s_1 10 \to 1 s_2 0$
$s_1 11 \to 1 s_2 1$	$s_1 1 \# \to 1 s_2 0 \#$	$s_2 00 \to 1 s_1 0$	$s_2 01 \to 1 s_1 1$
$s_2 0 \# \to 1 s_1 0 \#$	$0 s_2 1 \to s_3 01$	$1 s_2 1 \to s_3 11$	$\# s_2 1 \to \# s_3 01$
$0 s_3 0 \to s_4 00$	$1 s_3 0 \to s_4 10$	$\# s_3 0 \to \# s_4 00$	$0 s_3 1 \to s_2 00$
$1 s_3 1 \to s_2 10$	$\# s_3 1 \to \# s_2 00$	$0 s_4 0 \to s_{12} 01$	$1 s_4 0 \to s_{12} 11$
$\# s_4 0 \to \# s_{12} 01$	$0 s_4 1 \to s_9 00$	$1 s_4 1 \to s_9 10$	$\# s_4 1 \to \# s_9 00$
$s_5 00 \to 1 s_1 0$	$s_5 01 \to 1 s_1 1$	$s_5 0 \# \to 1 s_1 0 \#$	$0 s_5 1 \to s_6 00$
$1 s_5 1 \to s_6 10$	$\# s_5 1 \to \# s_6 00$	$0 s_6 0 \to s_7 00$	$1 s_6 0 \to s_7 10$
$\# s_6 0 \to \# s_7 00$	$0 s_6 1 \to s_7 01$	$1 s_6 1 \to s_7 11$	$\# s_6 1 \to \# s_7 01$
$0 s_7 0 \to s_8 00$	$1 s_7 0 \to s_8 10$	$\# s_7 0 \to \# s_8 00$	$0 s_7 1 \to s_6 00$
$1 s_7 1 \to s_6 10$	$\# s_7 1 \to \# s_6 00$	$0 s_8 0 \to s_7 00$	$1 s_8 0 \to s_7 10$
$\# s_8 0 \to \# s_7 00$	$s_8 10 \to 1 s_2 0$	$s_8 11 \to 1 s_2 1$	$s_8 1 \# \to 1 s_2 0 \#$
$s_9 00 \to 0 s_{19} 0$	$s_9 01 \to 0 s_{19} 1$	$s_9 0 \# \to 0 s_{19} 0 \#$	$0 s_9 1 \to s_4 01$
$1 s_9 1 \to s_4 11$	$\# s_9 1 \to \# s_4 01$	$0 s_{10} 0 \to s_4 01$	$1 s_{10} 0 \to s_4 11$
$\# s_{10} 0 \to \# s_4 01$	$s_{10} 10 \to 0 s_{13} 0$	$s_{10} 11 \to 0 s_{13} 1$	$s_{10} 1 \# \to 0 s_{13} 0 \#$
$0 s_{11} 0 \to s_4 00$	$1 s_{11} 0 \to s_4 10$	$\# s_{11} 0 \to \# s_4 00$	
$s_{12} 00 \to 0 s_{19} 0$	$s_{12} 01 \to 0 s_{19} 1$	$s_{12} 0 \# \to 0 s_{19} 0 \#$	$0 s_{12} 1 \to s_{14} 01$
$1 s_{12} 1 \to s_{14} 11$	$\# s_{12} 1 \to \# s_{14} 01$	$s_{13} 00 \to 0 s_{10} 0$	$s_{13} 01 \to 0 s_{10} 1$
$s_{13} 0 \# \to 0 s_{10} 0 \#$	$s_{13} 10 \to 1 s_{24} 0$	$s_{13} 11 \to 1 s_{24} 1$	$s_{13} 1 \# \to 1 s_{24} 0 \#$
$0 s_{14} 0 \to s_{15} 00$	$1 s_{14} 0 \to s_{15} 10$	$\# s_{14} 0 \to \# s_{15} 00$	$0 s_{14} 1 \to s_{11} 01$
$1 s_{14} 1 \to s_{11} 11$	$\# s_{14} 1 \to \# s_{11} 01$	$s_{15} 00 \to 0 s_{16} 0$	$s_{15} 01 \to 0 s_{16} 1$
$s_{15} 0 \# \to 0 s_{16} 0 \#$	$s_{15} 10 \to 1 s_{17} 0$	$s_{15} 11 \to 1 s_{17} 1$	$s_{15} 1 \# \to 1 s_{17} 0 \#$
$s_{16} 00 \to 0 s_{15} 0$	$s_{16} 01 \to 0 s_{15} 1$	$s_{16} 0 \# \to 0 s_{15} 0 \#$	$s_{16} 10 \to 1 s_{10} 0$
$s_{16} 11 \to 1 s_{10} 1$	$s_{16} 1 \# \to 1 s_{10} 0 \#$	$s_{17} 00 \to 0 s_{16} 0$	$s_{17} 01 \to 0 s_{16} 1$
$s_{17} 0 \# \to 0 s_{16} 0 \#$	$s_{17} 10 \to 1 s_{21} 0$	$s_{17} 11 \to 1 s_{21} 1$	$s_{17} 1 \# \to 1 s_{21} 0 \#$
$s_{18} 00 \to 0 s_{19} 0$	$s_{18} 01 \to 0 s_{19} 1$	$s_{18} 0 \# \to 0 s_{19} 0 \#$	$s_{18} 10 \to 1 s_{20} 0$
$s_{18} 11 \to 1 s_{20} 1$	$s_{18} 1 \# \to 1 s_{20} 0 \#$	$0 s_{19} 0 \to s_3 01$	$1 s_{19} 0 \to s_3 11$
$\# s_{19} 0 \to \# s_3 01$	$s_{19} 10 \to 1 s_{18} 0$	$s_{19} 11 \to 1 s_{18} 1$	$s_{19} 1 \# \to 1 s_{18} 0 \#$
$s_{20} 00 \to 1 s_{18} 0$	$s_{20} 01 \to 1 s_{18} 1$	$s_{20} 0 \# \to 1 s_{18} 0 \#$	$s_{20} 10 \to 0 s_{18} 0$
$s_{20} 11 \to 0 s_{18} 1$	$s_{20} 1 \# \to 0 s_{18} 0 \#$	$s_{21} 00 \to 0 s_{22} 0$	$s_{21} 01 \to 0 s_{22} 1$
$s_{21} 0 \# \to 0 s_{22} 0 \#$	$s_{21} 10 \to 1 s_{23} 0$	$s_{21} 11 \to 1 s_{23} 1$	$s_{21} 1 \# \to 1 s_{23} 0 \#$
$0 s_{22} 0 \to s_{10} 01$	$1 s_{22} 0 \to s_{10} 11$	$\# s_{22} 0 \to \# s_{10} 01$	$s_{22} 10 \to 1 s_{21} 0$
$s_{22} 11 \to 1 s_{21} 1$	$s_{22} 1 \# \to 1 s_{21} 0 \#$	$s_{23} 00 \to 1 s_{21} 0$	$s_{23} 01 \to 1 s_{21} 1$
$s_{23} 0 \# \to 1 s_{21} 0 \#$	$s_{23} 10 \to 0 s_{21} 0$	$s_{23} 11 \to 0 s_{21} 1$	$s_{23} 1 \# \to 0 s_{21} 0 \#$
$s_{24} 00 \to 0 s_{13} 0$	$s_{24} 01 \to 0 s_{13} 1$	$s_{24} 0 \# \to 0 s_{13} 0 \#$	$0 s_{24} 1 \to s_3 00$
$1 s_{24} 1 \to s_3 10$	$\# s_{24} 1 \to \# s_3 00 \ \}$		

We then use Theorem 2.2 to translate $G = (V_T, V_N, X_0, F_G)$ to a universal H system.
Let $V_H = \{0,1\} \cup \{T_\#\} \cup \{X_0, X_1, X_2\} \cup \{B\} \cup \{s_i \mid i \in 1..24\}$
Then the translation is :

$$
\begin{aligned}
V &= V_H \cup \{X, X', Y, Z\} \cup \{Y_\alpha \mid \alpha \in V_H\}\\
T &= \{0,1\}\\
A &= \{XBX_0Y, ZY, XZ,\\
&\quad ZT_\#s_1Y, ZT_\#Y, ZX_0Y, ZX_1Y, ZX_2Y,\\
&\quad ZY_\alpha, X'\alpha Z,\\
&\quad Z1s_10Y, Z1s_11Y, Z1s_10T_\#Y,\\
&\quad Zs_200Y, Zs_210Y, ZT_\#s_200Y, Z1s_20Y, Z1s_21Y, Z1s_20T_\#Y,\\
&\quad Zs_300Y, Zs_310Y, ZT_\#s_300Y, Zs_301Y, Zs_311Y, ZT_\#s_301Y,\\
&\quad Zs_400Y, Zs_410Y, ZT_\#s_400Y, Zs_401Y, Zs_411Y, ZT_\#s_401Y,\\
&\quad Z0s_50Y, Z0s_51Y, Z0s_50T_\#Y,\\
&\quad Zs_600Y, Zs_610Y, ZT_\#s_600Y,\\
&\quad Zs_700Y, Zs_710Y, ZT_\#s_700Y, Zs_701Y, Zs_711Y, ZT_\#s_701Y,\\
&\quad Zs_800Y, Zs_810Y, ZT_\#s_800Y,\\
&\quad Zs_900Y, Zs_910Y, ZT_\#s_900Y,\\
&\quad Zs_{10}01Y, Zs_{10}11Y, ZT_\#s_{10}01Y,\\
&\quad Z0s_{10}0Y, Z0s_{10}1Y, Z0s_{10}0T_\#Y, Z1s_{10}0Y, Z1s_{10}1Y, Z1s_{10}0T_\#Y,\\
&\quad Zs_{11}01Y, Zs_{11}11Y, ZT_\#s_{11}01Y,\\
&\quad Zs_{12}01Y, Zs_{12}11Y, ZT_\#s_{12}01Y,\\
&\quad Z0s_{13}0Y, Z0s_{13}1Y, Z0s_{13}0T_\#Y,\\
&\quad Zs_{14}01Y, Zs_{14}11Y, ZT_\#s_{14}01Y,\\
&\quad Zs_{15}00Y, Zs_{15}10Y, ZT_\#s_{15}00Y,\\
&\quad Z0s_{15}0Y, Z0s_{15}1Y, Z0s_{15}0T_\#Y,\\
&\quad Z0s_{16}0Y, Z0s_{16}1Y, Z0s_{16}0T_\#Y,\\
&\quad Z1s_{17}0Y, Z1s_{17}1Y, Z1s_{17}0T_\#Y,\\
&\quad Z0s_{18}0Y, Z0s_{18}1Y, Z0s_{18}0T_\#Y, Z1s_{18}0Y, Z1s_{18}1Y, Z1s_{18}0T_\#Y,\\
&\quad Z0s_{19}0Y, Z0s_{19}1Y, Z0s_{19}0T_\#Y,\\
&\quad Z1s_{20}0Y, Z1s_{20}1Y, Z1s_{20}0T_\#Y,\\
&\quad Z0s_{21}0Y, Z0s_{21}1Y, Z0s_{21}0T_\#Y, Z1s_{21}0Y, Z1s_{21}1Y, Z1s_{21}0T_\#Y,\\
&\quad Z0s_{22}0Y, Z0s_{22}1Y, Z0s_{22}0T_\#Y,\\
&\quad Z1s_{23}0Y, Z1s_{23}1Y, Z1s_{23}0T_\#Y,\\
&\quad Z1s_{24}0Y, Z1s_{24}1Y, Z1s_{24}0T_\#Y \mid \alpha \in V_H\ \}\\
R &= \{Xw\#s_100Y\$Z\#0s_50Y, \quad Xw\#s_101Y\$Z\#0s_51Y, \quad Xw\#s_10T_\#Y\$Z\#0s_50T_\#Y,\\
&\quad Xw\#s_110Y\$Z\#1s_20Y, \quad Xw\#s_111Y\$Z\#1s_21Y, \quad Xw\#s_11T_\#Y\$Z\#1s_20T_\#Y,\\
&\quad Xw\#s_200T\$Z\#1s_10Y, \quad Xw\#s_201T\$Z\#1s_11Y, \quad Xw\#s_20T_\#T\$Z\#1s_10T_\#Y,\\
&\quad Xw\#0s_21Y\$Z\#s_301Y, \quad Xw\#1s_21Y\$Z\#s_311Y, \quad Xw\#T_\#s_21Y\$Z\#T_\#s_301Y,\\
&\quad Xw\#0s_30Y\$Z\#s_400Y, \quad Xw\#1s_30Y\$Z\#s_410Y, \quad Xw\#T_\#s_30Y\$Z\#T_\#s_400Y,\\
&\quad Xw\#0s_31Y\$Z\#s_200Y, \quad Xw\#1s_31Y\$Z\#s_210Y, \quad Xw\#T_\#s_31Y\$Z\#T_\#s_200Y,\\
&\quad Xw\#0s_40Y\$Z\#s_{12}01Y, \quad Xw\#1s_40Y\$Z\#s_{12}11Y, \quad Xw\#T_\#s_40Y\$Z\#T_\#s_{12}01Y,\\
&\quad Xw\#0s_41T\$Z\#s_900Y, \quad Xw\#1s_41T\$Z\#s_910Y, \quad Xw\#T_\#s_41T\$Z\#T_\#s_900Y,\\
&\quad Xw\#s_500Y\$Z\#1s_10Y, \quad Xw\#s_501Y\$Z\#1s_11Y, \quad Xw\#s_50T_\#Y\$Z\#1s_10T_\#Y,\\
&\quad Xw\#0s_51Y\$Z\#s_600Y, \quad Xw\#1s_51Y\$Z\#s_610Y, \quad Xw\#T_\#s_51Y\$Z\#T_\#s_600Y,\\
&\quad Xw\#0s_60Y\$Z\#s_700Y, \quad Xw\#1s_60Y\$Z\#s_710Y, \quad Xw\#T_\#s_60Y\$Z\#T_\#s_700Y,\\
&\quad Xw\#0s_61Y\$Z\#s_701Y, \quad Xw\#1s_61Y\$Z\#s_711Y, \quad Xw\#T_\#s_61Y\$Z\#T_\#s_701Y,\\
&\quad Xw\#0s_70Y\$Z\#s_800Y, \quad Xw\#1s_70Y\$Z\#s_810Y, \quad Xw\#T_\#s_70Y\$Z\#T_\#s_800Y,
\end{aligned}
$$

$Xw\#0s_7 1Y\$Z\#s_6 00Y, \quad Xw\#1s_7 1Y\$Z\#s_6 10Y, \quad Xw\#T_\# s_7 1Y\$Z\#T_\# s_6 00Y,$
$Xw\#0s_8 0Y\$Z\#s_7 00Y, \quad Xw\#1s_8 0Y\$Z\#s_7 10Y, \quad Xw\#T_\# s_8 0Y\$Z\#T_\# s_7 00Y,$
$Xw\#s_8 10Y\$Z\#1s_2 0Y, \quad Xw\#s_8 11Y\$Z\#1s_2 1Y, \quad Xw\#s_8 1T_\# Y\$Z\#1s_2 0T_\# Y,$
$Xw\#s_9 00Y\$Z\#0s_{19} 0Y, \quad Xw\#s_9 01Y\$Z\#0s_{19} 1Y, \quad Xw\#s_9 0T_\# Y\$Z\#0s_{19} 0T_\# Y,$
$Xw\#0s_9 0Y\$Z\#s_4 01Y, \quad Xw\#1s_9 0Y\$Z\#s_4 11Y, \quad Xw\#T_\# s_9 0Y\$Z\#T_\# s_4 01Y,$
$Xw\#0s_{10} 0Y\$Z\#s_4 01Y, \quad Xw\#1s_{10} 0Y\$Z\#s_4 11Y, \quad Xw\#T_\# s_{10} 0Y\$Z\#T_\# s_4 01Y,$
$Xw\#s_{10} 10Y\$Z\#0s_{13} 0Y, \quad Xw\#s_{10} 11Y\$Z\#0s_{13} 1Y, \quad Xw\#s_{10} 1T_\# Y\$Z\#0s_{13} 0T_\# Y,$
$Xw\#0s_{11} 0Y\$Z\#s_4 00Y, \quad Xw\#1s_{11} 0Y\$Z\#s_4 10Y, \quad Xw\#T_\# s_{11} 0Y\$Z\#T_\# s_4 00Y,$
$Xw\#s_{12} 00Y\$Z\#0s_{19} 0Y, \quad Xw\#s_{12} 01Y\$Z\#0s_{19} 1Y, \quad Xw\#s_{12} 0T_\# Y\$Z\#0s_{19} 0T_\# Y,$
$Xw\#0s_{12} 1Y\$Z\#s_{14} 01Y, \quad Xw\#1s_{12} 1Y\$Z\#s_{14} 11Y, \quad Xw\#T_\# s_{12} 1Y\$Z\#T_\# s_{14} 01Y,$
$Xw\#s_{13} 00Y\$Z\#0s_{10} 0Y, \quad Xw\#s_{13} 01Y\$Z\#0s_{10} 1Y, \quad Xw\#s_{13} 0T_\# Y\$Z\#0s_{10} 0T_\# Y,$
$Xw\#s_{13} 10Y\$Z\#1s_{24} 0Y, \quad Xw\#s_{13} 11Y\$Z\#1s_{24} 1Y, \quad Xw\#s_{13} 1T_\# Y\$Z\#1s_{24} 0T_\# Y,$
$Xw\#0s_{14} 0Y\$Z\#s_{15} 00Y, \quad Xw\#1s_{14} 0Y\$Z\#s_{15} 10Y, \quad Xw\#T_\# s_{14} 0Y\$Z\#T_\# s_{15} 00Y,$
$Xw\#0s_{14} 1T\$Z\#s_{10} 01Y, \quad Xw\#1s_{14} 1T\$Z\#s_{10} 11Y, \quad Xw\#T_\# s_{14} 1T\$Z\#T_\# s_{10} 01Y,$
$Xw\#s_{15} 00Y\$Z\#0s_{16} 0Y, \quad Xw\#s_{15} 01Y\$Z\#0s_{16} 1Y, \quad Xw\#s_{15} 0T_\# Y\$Z\#0s_{16} 0T_\# Y,$
$Xw\#s_{15} 10Y\$Z\#1s_{17} 0Y, \quad Xw\#s_{15} 11Y\$Z\#1s_{17} 1Y, \quad Xw\#s_{15} 1T_\# Y\$Z\#1s_{17} 0T_\# Y,$
$Xw\#s_{16} 00Y\$Z\#0s_{15} 0Y, \quad Xw\#s_{16} 01Y\$Z\#0s_{15} 1Y, \quad Xw\#s_{16} 0T_\# Y\$Z\#0s_{15} 0T_\# Y,$
$Xw\#s_{16} 10Y\$Z\#1s_{10} 0Y, \quad Xw\#s_{16} 11Y\$Z\#1s_{10} 1Y, \quad Xw\#s_{16} 1T_\# Y\$Z\#1s_{10} 0T_\# Y,$
$Xw\#s_{17} 00Y\$Z\#0s_{16} 0Y, \quad Xw\#s_{17} 01Y\$Z\#0s_{16} 1Y, \quad Xw\#s_{17} 0T_\# Y\$Z\#0s_{16} 0T_\# Y,$
$Xw\#s_{17} 10Y\$Z\#1s_{21} 0Y, \quad Xw\#s_{17} 11Y\$Z\#1s_{21} 1Y, \quad Xw\#s_{17} 1T_\# Y\$Z\#1s_{21} 0T_\# Y,$
$Xw\#s_{18} 00Y\$Z\#0s_{19} 0Y, \quad Xw\#s_{18} 01Y\$Z\#0s_{19} 1Y, \quad Xw\#s_{18} 0T_\# Y\$Z\#0s_{19} 0T_\# Y,$
$Xw\#s_{18} 10Y\$Z\#1s_{20} 0Y, \quad Xw\#s_{18} 11Y\$Z\#1s_{20} 1Y, \quad Xw\#s_{18} 1T_\# Y\$Z\#1s_{20} 0T_\# Y,$
$Xw\#0s_{19} 0Y\$Z\#s_3 01Y, \quad Xw\#1s_{19} 0Y\$Z\#s_3 11Y, \quad Xw\#T_\# s_{19} 0Y\$Z\#T_\# s_3 01Y,$
$Xw\#s_{19} 10Y\$Z\#1s_{18} 0Y, \quad Xw\#s_{19} 11Y\$Z\#1s_{18} 1Y, \quad Xw\#s_{19} 1T_\# Y\$Z\#1s_{18} 0T_\# Y,$
$Xw\#s_{20} 00Y\$Z\#1s_{18} 0Y, \quad Xw\#s_{20} 01Y\$Z\#1s_{18} 1Y, \quad Xw\#s_{20} 0T_\# Y\$Z\#1s_{18} 0T_\# Y,$
$Xw\#s_{20} 10Y\$Z\#0s_{18} 0Y, \quad Xw\#s_{20} 11Y\$Z\#0s_{18} 1Y, \quad Xw\#s_{20} 1T_\# Y\$Z\#0s_{18} 0T_\# Y,$
$Xw\#s_{21} 00Y\$Z\#0s_{22} 0Y, \quad Xw\#s_{21} 01Y\$Z\#0s_{22} 1Y, \quad Xw\#s_{21} 0T_\# Y\$Z\#0s_{22} 0T_\# Y,$
$Xw\#s_{21} 10Y\$Z\#1s_{23} 0Y, \quad Xw\#s_{21} 11Y\$Z\#1s_{23} 1Y, \quad Xw\#s_{21} 1T_\# Y\$Z\#1s_{23} 0T_\# Y,$
$Xw\#0s_{22} 0Y\$Z\#s_{10} 01Y, \quad Xw\#1s_{22} 0Y\$Z\#s_{10} 11Y, \quad Xw\#T_\# s_{22} 0Y\$Z\#T_\# s_{10} 01Y,$
$Xw\#s_{22} 10Y\$Z\#1s_{21} 0Y, \quad Xw\#s_{22} 11Y\$Z\#1s_{21} 1Y, \quad Xw\#s_{22} 1T_\# Y\$Z\#1s_{21} 0T_\# Y,$
$Xw\#s_{23} 00Y\$Z\#1s_{21} 0Y, \quad Xw\#s_{23} 01Y\$Z\#1s_{21} 1Y, \quad Xw\#s_{23} 0T_\# Y\$Z\#1s_{21} 0T_\# Y,$
$Xw\#s_{23} 10Y\$Z\#0s_{21} 0Y, \quad Xw\#s_{23} 11Y\$Z\#0s_{21} 1Y, \quad Xw\#s_{23} 1T_\# Y\$Z\#0s_{21} 0T_\# Y,$
$Xw\#s_{24} 00Y\$Z\#0s_{13} 0Y, \quad Xw\#s_{24} 01Y\$Z\#0s_{13} 1Y, \quad Xw\#s_{24} 0T_\# Y\$Z\#0s_{13} 0T_\# Y,$
$Xw\#0s_{24} 1Y\$Z\#s_3 00Y, \quad Xw\#1s_{24} 1Y\$Z\#s_3 10Y, \quad Xw\#T_\# s_{24} 1Y\$Z\#T_\# s_3 00Y,$
$Xw\#Y\$Z\#T_\# s_1 Y,$
$Xw\#Y\$Z\#T_\# Y,$
$Xw\#X_1 0Y\$Z\#X_1 Y, \quad Xw\#0X_2 Y\$Z\#X_2 Y,$
$Xw\#X_1 1Y\$Z\#X_1 Y, \quad Xw\#1X_2 Y\$Z\#X_2 Y,$
$Xw\#X_1 T_\# Y\$Z\#X_2 Y, \quad Xw\#T_\# X_2 Y\$Z\#X_0 Y,$
$Xw\#s_{11} 1Y\$Z\#X_1 Y,$
$Xw\#s_i T_\# Y\$Z\#X_2 Y,$
$Xw\#\alpha Y\$Z\#Y_\alpha,$
$X'\alpha\#z\$X\#wY_\alpha,$
$X'w\#Y_\alpha\$Z\#Y,$
$X\#Z\$X'\#wY,$
$\#ZY\$XB\#xY,$
$\#Y\$XZ\# \quad \mid w \in V_H^*, \ i \in 1..24, \ x \in T^* \ \}$

The resulting extended H system $\gamma = (V, T, A, R)$ is then universal as it is the result of a transformation from a universal Turing machine. The proof of this is a direct result of composing the proof of Theorem 3.1 with the proofs of Theorem's 2.1 and 2.2. The latter two proofs are described in more detail in Salomaa [7] and Păun [3] respectively.
When one compares the complexity of the resulting H system with that of the original Turing machine, one obtains the following results :

$|V| = 2(n + m + 7)$
$|T| = n$
$|A| \leq (n + 1)(nm - |Final|) + 2(n + m + 5) + 8$
$|R| \leq (n + 1)(nm - |Final|) + 3(n + m + 5) + 2(n + 1) + m + 5 + |Final|$

where m is the number of states and n is the number of symbols of the Turing machine $\tau_r = (V_r, F_r)$, $Final = \{Xw\#saY\$Z\#X_1Y \in R \mid s \in \{s_i \mid i \in 1..24\},\ a \in T\}$, and $|R|$ is defined in respect to the number of groups of rules.
Note that while equality for $|R|$ is achievable simply through the use of an optimal Turing machine (where optimal implies that every state-symbol pair is used), equality for $|A|$ is not so simple. In fact if one assumes that every state is used then we may obtain a lower bound for $|A|$:

$$(n + 1)(m - 1) + 2(n + m + 5) + 8 \leq |A| \leq (n + 1)(nm - |Final|) + 2(n + m + 5) + 8$$

If we now consider the numerical values for the complexity of our universal extended H system we find that :

$|V| = 66$
$|T| = 2$
$|A| = 182$
$|R| = 270$

and thus we see that this agrees with our analytic results above with $|A| \in [139, 211]$ and with equality for $|R|$. We also note that $|A|$ will tend towards the lower bound when the respective Turing machine has a significant number of intensive states where a state is intensive iff it is the resultant state of more than two productions.

Acknowledgments

This research has been made possible through the efforts and contributions of many people. In particular I would like to thank my supervisor Prof. Cristian Calude for the assistance and opportunities that he has given me. Many thanks also go to Prof. Gheorghe Păun for numerous comments and suggestions made regarding this research.

References

[1] Erzsébet Csuhaj-Varjú, R. Freund, Lila Kari, and Gheorghe Păun. DNA computation based on splicing: Universality results. In *Biocomputing: Proceedings of the 1996 Pacific Symposium*, 1996.

[2] Tom Head. Formal language theory and DNA: an analysis of the generative capacity of specific recombinant behaviors. *Bulletin of Mathematical Biology*, 49(6):737–759, 1987.

[3] Gheorghe Păun. Regular extended H systems are computationally universal. *Journal of Automata, Languages, Combinatorics*, 1(1):27–36, 1996.

[4] Gheorghe Păun. Email to G. Alford, May 1997.

[5] Gheorghe Păun, Grzegorz Rozenberg, and Arto Salomaa. Computing by splicing. *Theoretical Computer Science*, 168:321–336, 1996.

[6] Yurii Rogozhin. Small universal turing machines. *Theoretical Computer Science*, 168: 215–240, 1996.

[7] Arto Salomaa. *Formal Languages.* Academic Press, New York, 1973.