Journal of Automata, Languages and Combinatorics **1** (1996) 1, 27–36

REGULAR EXTENDED H SYSTEMS ARE COMPUTATIONALLY UNIVERSAL [1]

GHEORGHE PĂUN
Institute of Mathematics of the Romanian Academy
P.O. Box 1–764, R–70700 Bucureşti, Romania
e-mail: gpaun@imar.ro

ABSTRACT

Recently, a new type of generative mechanisms were introduced, under the name of H systems. They are based on the splicing operation, a language-theoretic counterpart of DNA recombination. Extended H systems with finite sets of rules are known to generate regular languages only. If (at least) linear sets of rules are used, then characterizations of recursively enumerable languages are obtained. The power of the intermediate class of H systems, with regular sets of splicing rules, is not yet known. We settle here the question, by proving that extended H systems with finite sets of axioms and regular sets of rules characterize the recursively enumerable languages, thus having the full power of Turing machines (in fact, one axiom is shown to suffice).

Keywords: DNA recombination, splicing, H systems, Turing machines, descriptional complexity, DNA computing.

1. Introduction

The study of H systems is one of the newest and most promising branches of formal language theory. It has been originated by T. HEAD, by introducing the splicing operation [7], a language-theoretic model of DNA recombination. Based on this novel operation with strings and languages, generative mechanisms, called H systems, were defined. In the extended form considered in [12], such a system is a quadruple $\gamma = (V, T, A, R)$, where V is an alphabet, $T \subseteq V$ is the set of terminal symbols, $A \subseteq V^*$ is the set of axioms, and R is the set of splicing rules, namely strings of the form $r = u_1\#u_2\$u_3\#u_4$, with $u_1, u_2, u_3, u_4 \in V^*$ and $\#$, $\$$ two special symbols. Such a rule r is applied to x, y, yielding z, if $x = x_1u_1u_2x_2$, $y = y_1u_3u_4y_2$, $z = x_1u_1u_4y_2$. Applying iteratively the rules in R, starting from strings in A, and selecting only the strings composed of terminal symbols, we get a language, $L(\gamma)$. It is said that this is the language *generated* by γ.

The H systems can be classified according to the type of the languages A, R. In [2] it is proved that if A is regular and R finite, then $L(\gamma)$ is regular (no terminal alphabet is considered in [2], but the family of regular languages is closed under intersection,

[1]Research supported by the Academy of Finland, project 11281.