Splicing Systems and Automata Theory (Extended Abstract) *

Paola Bonizzoni †, Clelia De Felice**, Giancarlo Mauri †, Rosalba Zizza †

Introduction. Molecular Computing is a new branch connecting Biology, Computer Science and Mathematics: the idea is to look at DNA molecules as a model of computation where the operations are performed by using enzymes and other biological processes. The rapid increase in interest in Molecular Computing started with Adleman's article (L. M. Adleman (1994), Molecular computation of solutions to combinatorial problems, *Science* **226**, $1021-1024$), where the author showed how to solve (a small instance of) the Hamiltonian Path Problem, which is a NP-complete problem, by manipulating DNA sequences in a lab. The massive parallelism performed by molecular operations in the algorithm is the key element to its efficiency. Thanks to this interest in Molecular Computing, new models of computation have been suggested and old ones have been reconsidered. Indeed, in 1987 Tom Head pioneered a language-theoretic approach for studying some biological phenomena on DNA. He introduced the formalism of the *splicing system* to observe the generative power of some recombination processes. In this way, abstract models for the generated languages are defined as a formal counterpart of the DNA recombination under the reaction of restriction and ligase enzymes: the (linear and circular) Splicing Systems. A splicing operation is introduced as an operator on strings, inspired by a cut and paste molecular process. Formally, a splicing system is defined by giving an initial language I (initial set of DNA molecules) and a set of special words or rules R (enzymes). The set I is then transformed by repeated applications of the splicing operation. Intensive studies on linear splicing systems and variants have been done and the generating power of these systems has been compared with classical models in formal language theory. Our results deal with finite linear splicing systems, i.e., with both I, R finite sets. Looking for a characterization of the proper subclass of regular languages which are generated by these systems, we consider the still open problem of deciding whether or not a regular language $L \subseteq A^*$ belongs to this class of languages. Unlike the linear case, relatively few works on circular splicing systems and their languages have been published. A main result about the power of circular splicing systems is due to Pixton who proved that a circular regular language and a finite set of rules generate a regular language if the set of rules satisfies some additional hypotheses. We have focused our interest on what kind of regular circular languages are generated by circular splicing with both finite I and R and we give results for unary languages. We have also investigated the computational power of these systems and issues of algorithmic and descriptional complexity.

*Partially supported by 60% Project *"Linguaggi Formali e Modelli di Calcolo"* (University of Salerno) and by **MIUR** Project *"Linguaggi Formali e Automi: teoria ed applicazioni"*

†Dipartimento di Informatica Sistemistica e Comunicazione, Università degli Studi di Milano – Bicocca, Via Bicocca degli Arcimboldi 8, 20126 Milano – ITALY. E-mail:{bonizzoni,mauri,zizza}@disco.unimib.it ** Dipartimento di Informatica ed Applicazioni, Università di Salerno, 84081 Baronissi (SA) – ITALY. E-mail: defelice@unisa.it.