Journal of Automata, Languages and Combinatorics 5 (2000) 3, 255–268

AUTOMATICITY AND RATIONALITY[1]

JEFFREY SHALLIT[2]
Department of Computer Science, University of Waterloo
Waterloo, Ontario N2L 3G1 Canada
e-mail: shallit@graceland.uwaterloo.ca

ABSTRACT

Automaticity is a measure of descriptional complexity for formal languages L, and measures how closely L can be approximated by regular languages. I survey some of the known results and open problems on automaticity. I also discuss a measure which I call "rationality", and explain how it generalizes the well-known concept of linear complexity.

Keywords: Regular language, finite automata, formal power series, rational series, linear complexity, linear span, linear recurrence, linear feedback shift register sequence, automaticity, rationality, continued fraction.

1. Introduction

Let L be a formal language, that is, a subset of Σ^*, where Σ is a finite alphabet. We say L is *regular* if L is accepted by some finite automaton. Of course, not every language is regular, but we can approximate any language arbitrarily closely with regular languages. We say a language L' is an *n'th order approximation* to a language L if L and L' agree on all strings of length $\leq n$, that is, if $L \cap \Sigma^{\leq n} = L' \cap \Sigma^{\leq n}$, where by $\Sigma^{\leq n}$ we mean $(\Sigma \cup \{\epsilon\})^n$. The *automaticity function* $A_L(n)$ is defined to be the least number of states in any deterministic finite automaton (DFA) accepting an n'th order approximation to L.

Example 1 Let $|w|_a$ denote the number of occurrences of the letter a in the string w. Consider the language $L = \{w \in \{0,1\}^* \mid |w|_0 = |w|_1\}$. Then it can be shown that $A_L(n) = n + 1$ for $n \geq 0$.

Similarly, we define the *nondeterministic automaticity function* $N_L(n)$ to be the least number of states in any nondeterministic finite automaton (NFA) accepting an n'th order approximation to L.

---

[1]Full version of an invited lecture presented at the First International Workshop on *Descriptional Complexity of Automata, Grammars and Related Structures* held in Magdeburg, Germany, July 20 – 23, 1999.

[2]Supported in part by a grant from NSERC.