Journal of Automata, Languages and Combinatorics **13** (2008) 3/4, 269–297

AUTOMATA ACCEPTING BIFIX CODES

Zheng-Zhu Li
College of General Education, Aletheia University on Motou Campus
Matou, Taiwan 721
e-mail: lzz@mail.au.edu.tw

Huei-Jan Shyr
Department of Applied Mathematics, National Chung-Hsing University
Taichung, Taiwan 402
e-mail: hjshyr@amath.nchu.edu.tw

and

Yen-Shung Tsai
Department of Applied Mathematics, Chung-Yuan Christian University
Chung Li, Taiwan 320
e-mail: ystsai@math.cycu.edu.tw

ABSTRACT

Automata theory plays a very important role in the field of computer science. A language accepted by an automaton has been called a regular language in a standard way. Bifix codes are very important and useful codes in the whole code theory. The family of bifix codes have been divided into subfamilies such as comma-codes, comma-free codes and strict intercodes of index $m \geq 2$. In this paper we investigate the automata which accept variety of bifix codes. The language accepted by an automaton with one or two states can never be a prefix code or a suffix code. We obtained characterizations of an automaton with more than two states which accepted a prefix code and a suffix code. Many characterizations on automata with different number of states which accept different types of codes, such as bifix codes, infix codes, comma-codes and comma-free codes were nicely presented in this paper.

Keywords: Algorithm, finite automata, bifix codes, comma-free codes, comma-codes

1. Introduction and Definitions

In this paper we assume that the alphabet X contains more than one letter and let X^* be the free monoid generated by X. Every element of X^* is called a *word* or a *string* and let $X^+ = X^* \setminus \{1\}$, where 1 is the empty word. Every subset of X^* is a *language*. The cardinality of a language L is denoted by $|L|$. For $L \subseteq X^*$, as usual, we define the language L^* to be $L^* = \{1\} \cup L \cup L^2 \cup \cdots \cup L^n \cup \cdots$ and by definition, $\emptyset^* = \{1\}$.