Introduction

Automata theory arose as an interdisciplinary field, with roots in several scientific domains such as pure mathematics, electronics and computer science. This diversity appears in the material presented in this book which covers topics related to computer science, algebra, logic, topology and game theory.

The elementary theory of automata allows both the *specification* and the *verification* of simple properties of finite sequences of symbols. The possible practical applications include lexical analysis, text processing and sofware verification.

There are at least two possible extensions of this theory.

The theory of *formal series* is one of them. Words are replaced by functions associating to each word some numerical value. This value can be an integer counting the number of paths labeled by this word in an automaton or the integer represented by this word in some basis. It can also be a real number corresponding to some probability of the word.

The other possible extension is the subject of this book: Finite sequences of symbols are replaced by infinite sequences.

The motivation for this generalization originates in the early work of Richard Büchi in the sixties. Working on weak logical theories of the integers, he was lead to consider the monadic second order theory of the successor function on the integers. He was able to prove the decidability of this theory. He actually showed that all properties of the integers expressible in this logic can also be defined in terms of finite automata.

Later on, Robert McNaughton proved the equivalence of deterministic and non-deterministic automata, a natural extension of the corresponding result for finite words. This difficult result had been conjectured by David Muller while working at questions related to oscillating circuits.

Many other results have appeared since then and the theory has known an important increase of interest motivated by applications to problems in computer science. The notion of an infinite sequence is of interest to model the behavior of systems which are supposed to work endless, as operating systems for example.

This book presents a comprehensive treatment of all aspects of this theory. It gathers for the first time the basic results with the advanced ones. Actually, if several surveys have appeared on infinite words, this book is the first manual devoted to this topic. All proofs are given in detail, with a few, duly mentionned, exceptions.

The book is intended for researchers or advanced students in mathematics or computer science. No particular background is required to read it, except for a standard mathematical culture. The dependence between chapters is not too strong, making it possible to read some chapters independently from other ones.

The book can be used to lecture and the authors have used the manuscript for several years for graduate courses in computer science. It is unlikely to cover the whole content in one single course, but a selection with emphasis either on topology, or on logic, or on automata and semigroups, is possible.

The book is organized as follows. The first chapter contains the definitions of rational expressions, Büchi and Muller automata and recognizable sets. It covers the necessary elements of the theory of automata on finite words such as Kleene's theorem. A proof of McNaughton theorem is given, using Safra's determinization algorithm. Although this construction is rather involved, we have chosen to place it at the very beginning of the book because it is straightforward. Other proofs of McNaughton's theorem are given later on.

In the second chapter, we shift to a more algebraic point of view. The key idea of this chapter is to give a purely algebraic definition of recognizable sets. This point of view will be adopted quite often in the sequel. Our main tools are finite semigroups and their counterpart for infinite words, called ω-semigroups.

The third chapter introduces the topological aspects of the theory. It is really an excursion in the field known as descriptive set theory, situated at the border between analysis and logic. We show that the main notions introduced so far have a natural translation in terms of topology.

The fourth chapter is devoted to games. These games are two player mathematical games which are used as a tool to prove some results on infinite words and automata. For instance, in this chapter, games are used to prove the Büchi-Landweber theorem. Some particular games, such as Wadge games or Fraïssé-Ehrenfeucht games, are used in further chapters.

In the fifth chapter, we present a classification of recognizable sets of infinite words known as the Wagner hierarchy. It emphasizes once again the importance of finite semigroups in this theory.

Chapters VI and VII present the theory of varieties for infinite words. It is an extension of the so-called Eilenberg variety theory, which associates sets of finite words and families of finite semigroups. The families of semigroups are actually varieties of finite semigroups. The extension to infinite words leads to the notion of varieties of ω-semigroups. The classical result of Schützenberger on star-free sets and aperiodic semigroups founds its generalization with an appropriate notion of aperiodic ω-semigroups.

Logic comes in with Chapter VIII. The main point is that there is a close connexion between the concepts of automata theory and those of logic, as it was the case with topology. Thus, recognizability is equivalent with monadic second order definability while aperiodicity is equivalent to first order definability.

The last two chapters deal with two natural extensions of infinite words. The first one is concerned with two-sided infinite words, for which all notions generalize in a natural way. The second one deals with infinite trees. This case is important because of its role in the applications. The situation is very different with trees instead of words and, in particular, Büchi and Muller automata are no longer equivalent. The main result is Rabin's theorem which states the equivalence between recognizability by tree automata and monadic second order definability.

We wish to express our gratitude to all those who have helped us during the long maturation of this book. We thank Danièle Beauquier, Olivier Carton and

Wolfgang Thomas for their numerous suggestions on preliminary versions. We are also indebted to the numerous students and colleagues who have read our manuscripts and indicated errors of all sorts: André Arnold, Marie-Pierre Béal, Nicolas Bedon, Jean Berstel, Ahmed Bouabdallah, Mario Branco, Véronique Bruyère, Hugues Calbrix, Christian Choffrut, Joëlle Cohen, Pierre-Louis Curien, Volker Diekert, Paul Gastin, Dietrich Kuske, Bertrand Le Saec, Igor Litovski, Anca Muscholl, Bruno Petazzoni, Antoine Petit, Christophe Prieur, Roman R. Redziejowski, Jean-François Rey, Laurent Rosaz, Pierre-Yves Schobbens, Ludwig Staiger, Istvan Toth, Sébastien Veigneau, Klaus Wagner, Pascal Weil and Thomas Wilke.

The original manuscript was realized in AMS-LaTeX. The figures were realized using the packages pstricks and GasTeX.