Review of M. Lothaire, *Algebraic Combinatorics on Words*, Cambridge University Press, 2002.

By Jeffrey Shallit, School of Computer Science, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada.

Consider the following problem about strings of symbols. Call a string a *square* if consists of two adjacent copies of the same string. (An example is the English word `hotshots`, which is just $(\texttt{hots})^2$.) Must every sufficiently long string contain a square?

If the underlying alphabet is arbitrarily large, clearly the answer is no: for example, the string $\texttt{abcdef} \cdots \texttt{xyz}$ contains no square. On the other hand, if the alphabet consists of two symbols, the answer is yes: a simple argument shows that every string of length ≥ 4 contains a square. Over an alphabet of three symbols, however, the answer is no: consider the morphism μ such that $\mu(\texttt{a}) = \texttt{abc}$, $\mu(\texttt{b}) = \texttt{ac}$, and $\mu(\texttt{c}) = \texttt{b}$. If we iterate μ starting with the symbol $\texttt{a}$, then it can be shown that each of the strings $\mu(\texttt{a}) = \texttt{abc}$, $\mu^2(\texttt{a}) = \mu(\mu(\texttt{a})) = \texttt{abcacb}$, etc., contains no square.

This beautiful result is due to the Norwegian mathematician Axel Thue in 1906, and is the first significant result in what is now called *combinatorics on words* — *word* being another term for a string of symbols. In the past thirty years this field has seen explosive growth, particularly in France and Italy.

The book under review is a collection of 13 chapters written by different authors, each chapter dealing with some aspect of combinatorics on words. (The name of the author, M. Lothaire, is a collective pseudonym derived from Lotharingia, an ancient name for the area comprising roughly modern-day Lorraine, Alsace, Belgium, Luxembourg, and north-west Germany.) The authors and titles are as follows: Jean Berstel and Dominique Perrin, "Finite and infinite words"; Jean Berstel and Patrice Séébold, "Sturmian words"; Julien Cassaigne, "Unavoidable patterns"; Aldo de Luca and Stefano Varricchio, "Sesquipowers"; Alain Lascoux, Bernard Leclerc, and Jean-Yves Thibon, "The plactic monoid"; Véronique Bruyère, "Codes"; Christiane Frougny, "Numeration systems"; Filippo Mignosi and Antonio Restivo, "Periodicity"; Christophe Reutenauer, "Centralizers of noncommutative series and polynomials"; Dominique Foata and Guo-Niu Han, "Transformations on words and q-calculus"; Jacques Désarménien, "Statistics on permutations and words"; Volker Diekert, "Makanin's algorithm"; and Tero Harju, Juhani Karhumäki, and Wojciech Plandowski, "Independent systems of equations".

Many of the chapters present material that is difficult to find or unavailable elsewhere. Chapter 2 on Sturmian words is particulary noteworthy. Sturmian words can be defined in many equivalent ways, but the simplest is probably as follows: an infinite word over a two-letter alphabet is Sturmian if it has exactly $n+1$ distinct subwords (factors) of length n, for all $n \geq 0$. An example is the infinite Fibonacci word

$$\texttt{0100101001001} \cdots ,$$

which is the fixed point of the morphism sending $\texttt{0} \rightarrow \texttt{01}$ and $\texttt{1} \rightarrow \texttt{0}$. Sturmian words have many interesting combinatorial and number-theoretic properties, and have arisen in fields as diverse as computer graphics (where they arise in the display of digitized straight lines), symbolic dynamics, and music. This comprehensive chapter collects many results from the literature on Sturmian words, sometimes with new proofs.

Chapter 7 discusses numeration systems, that is, generalizations of the familiar representation of numbers in base 10. The author covers β-expansions, which are representations in an arbitrary real base $\beta > 1$ and the interesting connections between these expansions and Pisot and Perron numbers. Also discussed is the representation of complex numbers in Gaussian integer bases and the connection to fractals.

Chapter 8 is about periodicity. Here one of the most famous results is the Fine-Wilf theorem, which states that if two periodic infinite words with periods p and q agree on $p + q - \gcd(p, q)$ consecutive symbols, then they are identical. Recently this theorem has been generalized by several different groups of researchers; among other things, this chapter discusses a generalization to three periods.

Although none of the authors is a native English speaker, the exposition is generally quite good. Sometimes a Gallicism creeps in, as on p. 165, where the crucial sentence describing the main topic of the chapter is hard to parse because of poor agreement. I did notice a number of minor errors: on p. 20, for example, a transducer is said to process an input beginning with the most significant digit, but actually begins with the least significant digit. And, as inevitably occurs with chapters written by different authors, there is sometimes a clash of notation: x^ζ denotes a set of words on p. 12, and a word on p. 34, but ζ itself denotes a word on p. 115. But overall the material is presented comprehensibly, even pleasingly. Each chapter is accompanied by exercises and detailed notes on the origins of the subject. A bibliography with about 500 references makes it easy to pursue further research.

The volume under review is actually the second by the same pseudonymous author; the first is entitled *Combinatorics on Words* and a third volume is currently being written. However, this book can be read largely independently fo the first volume, and will be of interest to number theorists, combinatorialists, theoretical computer scientists, and algebraists.