Codes and Automata
Corrections and Complements

January 14, 2016

This file contains corrections and complements to the book.

1 Preliminaries

- p. 28 ℓ. -2 : Insert 'provided the automaton is complete' after 'The matrix M/k is stochastic'.
- p. 30 ℓ. Replace lines 3–17 by :

  Applying by induction the theorem to U and W, we obtain nonnegative eigenvectors u and w for the eigenvalues ρ_U and ρ_W of U and W. We prove that $\max(\rho_U, \rho_W)$ is an eigenvalue of M with some nonnegative eigenvector.

  If $\rho_U \geq \rho_W$, then ρ_U is an eigenvalue of M with the corresponding eigenvector $\left[\begin{smallmatrix} u \\ 0 \end{smallmatrix}\right]$. If $\rho_U < \rho_W$, then we show that ρ_W is an eigenvalue of M for the eigenvector $\left[\begin{smallmatrix} u' \\ w \end{smallmatrix}\right]$, where

  $$u' = \Big(\sum_{n \geq 0} U^n \rho_W^{-n-1}\Big) V w = (\rho_W I - U)^{-1} V w\,.$$

  Since $\rho_U < \rho_W$, the series $\sum_{n\geq 0} U^n \rho_W^{-n}$ converges in view of Proposition 1.9.3, and it converges to a matrix with nonnegative coefficients because each U^n has nonnegative coefficients. If follows that u' has nonnegative coefficients. Moreover

  $$Vw = (\rho_W I - U)u' = \rho_W u' - U u'\,,$$

  showing that $M\left[\begin{smallmatrix} u' \\ w \end{smallmatrix}\right] = \rho_W \left[\begin{smallmatrix} u' \\ w \end{smallmatrix}\right]$. This shows that $\rho_M \geq \max(\rho_U, \rho_W)$. Conversely, if λ is an eigenvalue of M with corresponding eigenvector $\left[\begin{smallmatrix} u \\ v \end{smallmatrix}\right]$, then λ is an eigenvalue of W if $v \neq 0$, and is an eigenvalue of U if $v = 0$. This proves that $\rho_M = \max(\rho_U, \rho_W)$.
- p. 31 ℓ. 12–13 replace by: Recall that the adjacency matrix of a complete deterministic automaton over a k-letter alphabet has spectral radius k and...
- p. 37 ℓ. 2 of proof of Proposition 1.10.10 : remove the last '$\times$'.

2 Codes

- p. 74 ℓ. 16 : Insert '$= pqt^2 F_{D_a^*}(t)$' after '$F_a(t)F_{D_a^*}(t)F_b(t)$'
- p. 102 ℓ. 3 of Exercise 2.4.2 : Replace 'prefix of w' by 'prefix u of w'
- p. 102 In Exercise 2.4.3, replace the second sentence by: Let $D = D_n$ be the Dyck code on A (Example 2.2.12). Show that one has

$$\begin{aligned} f_D(t) &= \frac{n}{2n-1}(1 - \sqrt{1 - 4(2n-1)t^2}), \\ f_{D^*}(t) &= \frac{1 - n + n\sqrt{1 - 4(2n-1)t^2}}{1 - 4n^2t^2}. \end{aligned}$$

3 Prefix codes

- p. 114 Figure 3.8(b) : Replace the label 'a' by 'b' on the last edge of the path of length 3.
- p. 117 ℓ. -8 : Replace 'minimal automata' by 'minimal automaton'
- p. 157 ℓ. -7 : Replace '$\mathcal{B}$' by '$\mathcal{B}$ of the proof of Lemma 3.8.6'
- p. 173 Exercise 3.8.2 ℓ. 1 : Add 's' to 'length'
- p. 173 Exercise 3.8.2 ℓ. 3 : Insert '3.8.1 and' before '3.6.4'
- p. 173 add the following exercise, due to Staiger (2007). It shows that for a any infinite prefix code, there is a maximal prefix code on the same alphabet which has the same length distribution.
  Exercise 3.8.3 Let X be an infinite prefix code. Let $x_1, x_2, \ldots$ be an enumeration of X by nondecreasing lengths. Set $\ell_n = |x_n|$. Let $X_1 \subset X_2 \subset \cdots$ be the strictly increasing sequence of prefix codes defined as follows.
  Set $X_1 = \emptyset$. Assume that X_n is already defined and define X_{n+1} as follows. Set $m = \mathrm{Card}(X_n)$ and $\ell = \ell_{m+1}$. Let $\{u_1, \ldots, u_t\}$ be the set of words of length ℓ without any prefix in X_n. For $1 \le i \le t$, let v_i be a word such that $u_i v_i$ has length ℓ_{m+i}. Then $X_{n+1} = X_n \cup \{u_1 v_1, \ldots, u_t v_t\}$.
  Let X' be the union of the X_n. Show that:
  1. the length distribution of X and X' are the same.
  2. the set X' is a maximal prefix code.

4 Automata

- p. 194 Example 4.3.5 : 'the code $X =$' instead of 'the code $C =$'
- The *profinite metric* on a monoid M is the topology induced by the distance $d(u, v) = 2^{-n}$ where n is the minimal cardinality of a monoid N for which there is a morphism $\varphi : M \to N$ such that $\varphi(u) \neq \varphi(v)$. The free profinite monoid on A, denoted $\widehat{A^*}$, is the completion of the free monoid A^* for the profinite metric (see Almeida (1994)). It is a topological monoid, that is, a monoid with a topology for which the multiplication is continuous.
  The aim of this exercise (taken from Margolis et al. (1998)) is to expore the notion of a code in the free profinite monoid. Any morphism $\beta : B^* \to A^*$ extends uniquely by continuity to a continuous morphism $\hat{\beta} : \widehat{B^*} \to \widehat{A^*}$. A set $X \subset \widehat{A^*}$ is called a *profinite code* if the continuous extension $\hat{\beta}$ of any bijection $\beta : B \to X$ is injective.

Exercise 4.3.1 Show that any finite code $X \subset A^+$ is a profinite code.
Solution: Let $\beta : B^* \to A^*$ be a coding morphism for X. We have to show that for any pair $u, v \in \widehat{B^*}$ of distinct elements, we have $\hat{\beta}(u) \neq \hat{\beta}(v)$, that is, there is a continuous morphism $\hat{\alpha} : \widehat{A^*} \to M$ into a finite monoid M such that $\hat{\alpha}\hat{\beta}(u) \neq \hat{\alpha}\hat{\beta}(v)$. For this, let $\psi : \widehat{B^*} \to N$ be a continuous morphism into a finite monoid N such that $\psi(u) \neq \psi(v)$. Let P be the set of proper prefixes of X and let $\mathcal{T}$ be the prefix transducer associated to β. Let α be the morphism from A^* into the monoid of $P \times P$-matrices with elements in $N \cup 0$ defined as follows. For $x \in A^*$ and $p, q \in P$, we have

$$\alpha(x)_{p,q} = \begin{cases} \psi(y) & \text{if there is a path } p \xrightarrow{x|y} q \\ 0 & \text{otherwise.} \end{cases}$$

Then $M = \alpha(A^*)$ is a finite monoid and α extends to a continuous morphism $\hat{\alpha} : \widehat{A^*} \to M$. Since, by Proposition 4.3.2, the transducer $\mathcal{T}$ realizes the decoding function of X, we have $\alpha\beta(y)_{1,1} = \psi(y)$ for any $y \in B^*$. By continuity, we have $\hat{\alpha}\hat{\beta}(y)_{1,1} = \psi(y)$ for any $y \in \widehat{B^*}$. Then $\hat{\alpha}$ is such that $\hat{\alpha}\hat{\beta}(u) \neq \hat{\alpha}\hat{\beta}(v)$. Indeed $\hat{\alpha}\hat{\beta}(u)_{1,1} = \psi(u) \neq \psi(v) = \hat{\alpha}\hat{\beta}(v)_{1,1}$.

5 Deciphering delay

- p. 214 ℓ. 15 : Insert 'with $a \in A$' at the beginning of the line
- p. 221 Add the following exercise which is a result from Simon (1990).
  Exercise. A *rectangular band* is a semigroup of the form $I \times \Lambda$ for two sets I, Λ with the multiplication
  $$(i, \lambda)(j, \mu) = (i, \mu)$$
  for $i, j \in I$ and $\lambda, \mu \in \Lambda$.
  Let $f : A^+ \to S$ be a morphism from A^+ onto a rectangular band. Show that for any $s \in S$, the semigroup $f^{-1}(s)$ is of the form X^+ where X is a code with verbal dechiphering 1.
  Solution. Assume that $xyu = x'y'$ with $x, x', y \in X$, $y' \in X^*$ and $u \in A^*$. Assume that $x = x'v$. Then $y' = vyu$ implies $f(v)\mathcal{R}f(y') = s$ and $x = x'v$ implies $f(v)\mathcal{L}f(x) = s$. Thus $f(v) = s$ which implies $v \in X^*$. This shows that $x = x'$.

6 Bifix codes

- p. 227 ℓ. 11 : 'Proposition' instead of 'Theorem'
- p. 229 ℓ. 9 : 'any parse of v' instead of 'any parse of u'
- p. 230 ℓ. 2 : Replace 'Theorem 3.1.6' by 'Proposition 3.1.3', and insert 'by Proposition 3.1.6' before '$1 - \underline{X}$'.
- p. 233 ℓ. 5 : 'for $k = 0, 1$' instead of 'for $k = 0, 1, 2$'
- p. 234 ℓ. -8 : 'Corollary' instead of 'Proposition'
- p. 245 ℓ. 2 of Proposition 6.3.14 : '$H = A^- X A^-$' instead of '$H = A^* \setminus X A^-$
- p. 274 ℓ. 7 : Insert 'Exercise 6.1.2 is from Reutenauer (1979)'

7 Circular codes

- p. 291 line -15 change X_3 to $X_3 = \{ab, aab, bab, aaab, baab, bbab, \ldots\}$.
- p. 297 Add the following exercises for Section 7.1.

  Exercise 7.1.3

  Let B_n be an alphabet with n elements and let $\bar{B_n} = \{\bar{b} \mid b \in B_n\}$. Let $A_n = B_n \cup \bar{B_n}$. Consider the congruence $\equiv$ of A_n^* generated by all the relations $b\bar{b} \equiv 1$ for $b \in B_n$. Let M be the corresponding quotient monoid and let $\varphi : A^* \to M$ be the corresponding morphism. The set $\varphi^{-1}(1)$ is a free submonoid generated by a bifix code D'_n called the *restricted Dyck code*. Let $R = A_n^* \setminus A_n^*\{b\bar{b} \mid b \in B_n\}A_n^*$. Show that R is a set of representatives of the classes modulo $\equiv$.

  Identify M and R. Show that an element $w \in R$ is right-invertible (resp. left-invertible) if and only if $w \in B_n^*$ (resp. $w \in \bar{B_n}^*$). Deduce that if $uv, vu \in D_n'^*$, then $u, v \in D_n'^*$. Conclude that D'_n is a circular code.

  Solution. By induction on the length of $u \in R$. If $uv \equiv 1$ for some $v \in R$, we have $u = u'b$ and $v = \bar{b}v'$ with $b \in B$ and $u'v' \equiv 1$. By induction $u' \in B^*$. Thus $u \in B^*$.

  Exercise 7.1.4 Let D'_n be the restricted Dyck code as above. Show that one has the following disjoint union.

$$D_n'^* \setminus \{1\} = \bigcup_{b \in B} bD_n'^*\bar{b}D_n'^*.$$

  Let $g_n(t)$ (resp. $h_n(t)$) be the generating series of $D_n'^*$ (resp. D'_n). Show that $g_n(t) = (1 - h_n(t))^{-1}$ and that $g_n(t) = 1 + nt^2 g_n(t)^2$. Deduce that $g_n(t) = (1 - \sqrt{1 - 4nt^2})/2nt^2$ and that $h_n(t) = (1 - \sqrt{1 - 4nt^2})/2$. Note that the value $h_1(t) = (1 - \sqrt{1 - 4t^2})/2$ is consistent with the value given for $F_{D_a}(t) = h_1(t/2)$ for $p = q = 1/2$ in Example 2.4.10.

  Using the binomial formula, as in the derivation of Equation (3.13), show that $g_n(t) = \sum_{k \geq 0} C_k n^k t^{2k}$, where $C_k = \frac{1}{k+1}\binom{2k}{k}$ is the k-th Catalan number (see Table 3.1 p. 129). Thus

$$\begin{aligned} g_1(t) &= 1 + t^2 + 2t^4 + 5t^6 + 14t^8 + 42t^{10} + 132t^{12} + 429t^{14} + \ldots, \\ g_2(t) &= 1 + 2t^2 + 8t^4 + 40t^6 + 224t^8 + 1344t^{10} + 8448t^{12} + \ldots \end{aligned}$$

  In particular, $g_1(t) = \sum_{k \geq 0} C_k t^{2k}$ and C_k is the number of words of length $2k$ in $D_1'^*$. Give a direct bijection between the set of words of length $2k$ in $D_n'^*$ and the Cartesian product of the set of words of length $2k$ in $D_1'^*$ with B_n^k.

  Solution. The words in $D_n'^*$ may classically viewed as well-parenthesized expressions, with n different types of parenthesis; each such word, of length $2k$, defines a unique word of length $2k$ in $D_1'^*$, by matching the opening and closing parenthesis; the sequence of length k of opening parenthesis, from left to right, defines a word of length k in B_n^*. This gives the desired bijection. The various Dyck and restricted Dyck codes are described in more detail in (Berstel,1979).

- p. 298 Add the following exercise for Section 7.3 (see Stanley, 1997).
Exercise 7.3.6 Let X be a circular code. For $x \in X$ and $n \geq 0$, let $g_{x,n}$ be the number of words of length n having an interpretation (s, y, p) with $x = ps$ and p nonempty. Show that

$$g_{x,n} = |x| \operatorname{Card}(X^* \cap A^{n-|x|}) \tag{1}$$

Deduce from this equality a direct proof of Equation (7.14).
Solution. Let S be the set of words having a conjugate in X^*. Set $u_n^* = \operatorname{Card}(X^* \cap A^n)$ and $u^*(z) = \sum_{n \geq 0} u_n z^n$. Since X is circular, any word in S has a unique interpretation (s, y, p) such that $ps \in X$ and p nonempty. Thus $g_{x,n} = |x| u_{n-|x|}^*$ which proves (1). Since $p_n = \sum_{x \in X} g_{n,x}$, we obtain

$$\begin{aligned} p_n &= \sum_{x \in X} g_{n,x} = \sum_{x \in X} |x| u_{n-|x|}^* \\ &= \sum_{m=0}^{n} m u_m u_{n-m}^*. \end{aligned}$$

This shows that $p(z) = zu'(z)u^*(z)$ whence Formula (7.14).
- Add to the Notes: The Hall sequences defined in Section 3 are named after Hall (1934) (they are actually related to the Lazard sets as defined in Chapter 8, see Proposition 0.1 below and the Notes of Chapter 8 below).
- p. 299 Add after Theorem 7.3.7 is due to Schützenberger: It was conjectured by Golomb and Gordon (1965) and obtained inpendently by Scholtz (1969).

8 Factorizations of free monoids

- Add the following proposition before Example 8.1.8.
The following result connects Hall sequences and Lazard sets.

PROPOSITION 0.1 *Let $(x_n)_{n \geq 1}$ be a Hall sequence and let $(X_n)_{n \geq 1}$ be the associated sequence of codes. The set $Z = \{x_n \mid n \geq 1\}$ with the order defined by the indices is a Lazard set if and only if for every $n \geq 1$ there is some $j \geq 1$ such that $X_j \cap A^{[n]} = \emptyset$.*

Proof The condition is clearly necessary. Conversely, if the Hall sequence satisfies this condition, let $n \geq 1$ and let $j \geq 1$ be such that $X_j \cap A^{[n]} = \emptyset$. Let $Z \cap A^{[n]} = \{z_1, z_2, \ldots, z_k\}$ with $z_1 < z_2 < \cdots < z_k$ and let $Z_1, \ldots, Z_k$ be the sets defined by $Z_1 = A$ and $Z_{i+1} = z_i^*(Z_i \setminus z_i)$ with $z_i \in Z_i$ for $1 \leq i \leq k$. Then, we have $z_1 = x_{i_1}, \ldots, z_k = x_{i_k}$ with $i_1 < \cdots < i_k$. Assume that there is a word $z \in Z_{k+1}$ of length at most n. Since $X_j \cap A^{[n]} = \emptyset$, there is some ℓ with $i_k < \ell < j$ such that $z = x_\ell$. But this contradicts the definition of k. Thus $Z_{k+1} \cap A^{[n]} = \emptyset$. ■

- p. 323 line -5 change $L \cap A^n$ into $L \cap A^{[n]}$.
- Add the following exercises.

Exercise 8.2.11 The aim of this exercise is to generalize the notion of bisection. Let F be a factorial set. A *bisection* of F is a pair (X, Y) of subsets of F such that $\underline{F} = \underline{X}^*\underline{Y}^*$.

(i) Show that

$$Y^*X^* \cap F \subset X^* \cup Y^*$$

(ii) Show that X is $(1, 0)$-limited and Y is $(0, 1)$-limited.

Solution: (i) It is enough to show that $YX \cap F \subset X \cup Y$. From $\underline{F} = \underline{X}^*\underline{Y}^*$, we deduce $\underline{F}^{-1} = 1 - \underline{X} - \underline{Y} + \underline{Y}\,\underline{X}$. Since F is factorial, we have $(\underline{F}^{-1}, w) = 0$ for any word $w \in F$ of length at least 2. This implies the conclusion.

(ii) Assume that $uv \in X^*$. Then $u, v \in F$ implies $u = xy$, $v = x'y'$ for some $x, x' \in X^*$ and $y, y' \in Y^*$. By (i), we have $yx' \in X^* \cup Y^*$. By uniqueness of the factorization, we have $yx' \in X^*$ and $y' = 1$. Thus $v \in X^*$.

The following exercise is from Keller (1991) and Béal and Dima (2015)

Exercise 8.2.12 Let D be the one-sided Dyck code on the alphabet $A \cup \bar{A}$. It is the class of 1 for the congruence generated by the relations $a\bar{a} = 1$ for $a \in A$. Let F be the set factors of D.

(i) Show that $(D^*\bar{A}, D \cup A)$ is a bisection of F.

(Hint: show that a reduced word with repect to the rules rewriting any $a\bar{a}$ into 1 for $a \in A$ is in $A^*\bar{A}^*$.

(ii) Let $f(t)$ be the generating series of F. Show that

$$f(t) = \frac{1 + \sqrt{1 - 4nt^2}}{(1 - 2nt + \sqrt{1 - 4nt^2})^2}$$

with $n = \mathrm{Card}(A)$.

(iii) Show that the radius of convergence of the generating series of F is $\frac{1}{n+1}$.

Solution: (i) Let $\varphi : (A \cup \bar{A})^* \to \mathbb{Z}$ be the morphism defined by $\varphi(a) = 1$ if $a \in A$ and $\varphi(a) = -1$ if $a \in \bar{A}$. For $w \in F$, set $w = uv$ where u is the shortest prefix of w such that $\varphi(u') \geq \varphi(u)$ for any prefix of u. Then $u \in (D^*\bar{A})^*$ and $v \in (D \cup A)^*$.

(ii) Let $g(t), h(t)$ be the generating series of $D^*\bar{A}, D \cup A$. By Exercise 2, we have $g(t) = (1 - \sqrt{1 - 4t^2})/2nt^2$ and $h(t) = (1 - \sqrt{1 - 4nt^2})/2$. Since, by (i),

$$f(t) = \frac{1}{(1 - ntg(t))(1 - nt - h(t))} = \frac{1 - h(t)}{(1 - nt - h(t))^2}$$

the result follows.

(iii) The value $\rho = \frac{1}{n+1}$ is solution of $1 - n\rho - h(\rho) = 0$ and is thus a pole of f. The other singlarity of f is the value $t = 1/2\sqrt{n}$ for which $\sqrt{1 - 4nt^2} = 0$. For $n \geq 1$, we have $\frac{1}{n+1} \leq \frac{1}{2\sqrt{n}}$ and thus $\frac{1}{n+1}$ is the radius of convergence of f.

Exercise 8.2.13 Show that $D \cup A$ is a circular code (this implies that D itself is a circular code, see Exercise 7.1.3 in this fascicule).
Solution: This follows from Exercise 8.2.12 and 8.2.11.

Execise 8.2.14 Show that any factor of D^* has a conjugate in $(D^*\bar{A})^*$ or in $(D \cup A)^*$.
Solution: Set $X = D^*\bar{A}$ and $Y = D \cup A$. By Exercise 8.2.12, the pair (X, Y) is a bisection of the set F of factors of D. Thus the statement follows from Exercise 8.2.11 (i).

- Add to the Notes, line -14: Lazard sets, also called *Hall sets*, are used to build bases of free Lie algebras (see Lothaire (1997), Viennot (1978), Reutenauer (1993),Bokut and Chibrikov (2006)). The bracketting of a word z in a Lazard set Z is defined by $z \mapsto (x, y)$ where $z = xy$ and x is the longest proper prefix of z which is in Z (see Lothaire (1997)). The expressions obtained can be used to define ring (or Lie) commutators $xy - yx$ or group commutators $xyx^{-1}y^{-1}$. The term 'Lazard set' is by reference to a method called *Lazard elimination method*, Lazard (1954). The term *Hall set* is by reference to an algorithm of Hall (1934) called the *collecting process* in Hall (1976).

10 Synchronization

- p. 395 Section 10.6 Notes : Insert 'The notion of *constant* appears in Schützenberger (1975). The notion of *synchronizing word* appears in many contexts with various denominations, including magic word (Lind, Marcus (1995)) or reset sequence. It has been defined in Chapter 3 for prefix codes and for deterministic automata. The notion of *synchronizing pair* is an extension of the definition of synchronizing word to codes which are not prefix. It is due to Schützenberger (1979b).'
- p. 395 Section 10.6 Notes ℓ. 3 : Insert before 'However' the sentence 'This is Theorem 10.2.11.'

11 Groups of codes

- p. 401 Delete 'It is not known ...thin maximal codes'.
- p. 412 ℓ 5 Replace 'Example 3.6.6' by 'Example 3.6.3'
- p. 415 Remark 11.4.5, ℓ. 3 : Insert a space between 'X' and 'is'
- p. 433 It has been shown by Yun Liu (2012) that the generalization of Proposition 11.1.6 for a code which is not prefix is false. Proposition 11.2.3 already appears as Property 2 in Schützenberger (1964) with the hypothesis that $G(X)$ is abelian. It has been shown in Liu (2012) that the corresponding statement for a code which is not prefix is false.

13 Densities

- p. 452, ℓ. 1 : Replace the first paragraph by:

A real valued function μ defined on a Boolean algebra of sets $\mathcal{F}$ is *additive* if for any disjoint sets $E, F \in \mathcal{F}$, one has $\mu(E \cup F) = \mu(E) + \mu(F)$. It is called *countably additive* if

$$\mu(\bigcup_{n\geq 0} E_n) = \sum_{n\geq 0} \mu(E_n)$$

for any sequence $(E_n)_{n\geq 0}$ of pairwise disjoint sets in $\mathcal{F}$ such that $\bigcup_{n\geq 0} E_n \in \mathcal{F}$. If μ is additive and takes nonnegative values, then it is *monotone* in the sense that if $E \subset F$ for $E, F \in \mathcal{F}$, then $\mu(E) \leq \mu(F)$ since indeed $\mu(F) = \mu(E \cup (F \setminus E)) = \mu(E) + \mu(F \setminus E) \geq \mu(E)$.

- p. 452, ℓ. 8 : Replace Proposition 13.1.3 by:

  Let μ be a countably additive function defined on a Boolean algebra $\mathcal{F}$ of sets. Then

  $$\mu(\bigcup_{n\geq 0} E_n) \leq \sum_{n\geq 0} \mu(E_n)$$

  for any sequence $(E_n)_{n\geq 0}$ of sets in $\mathcal{F}$ such that $\bigcup_{n\geq 0} E_n \in \mathcal{F}$.
- p. 456 : Replace Proposition 13.1.13 by:

  The function μ satisfies $\mu(A^\omega) = 1$ and is countably additive.
- p. 457 ℓ. 8 : Add 'The second inequality holds by Proposition 13.1.3 since, by Lemma 13.1.12, $\mathcal{F}$ is a Boolean algebra.'
- p. 457 ℓ. 10 : Replace the sentence by: A function ν defined on a family of sets $\mathcal{F}$ is called *countably subadditive* if for any sequence $(E_n)_{n\geq 0}$ of sets in $\mathcal{F}$ such that $\bigcup_{n\geq 0} E_n \in \mathcal{F}$, one has $\nu(\bigcup_{n\geq 0} E_n) \leq \sum_{n\geq 0} \nu(E_n)$.
- p. 459 ℓ. 15 : Replace 'Then Equation (13.1) holds' by 'Then the equation of line 6 holds'
- p. 495 proof of Proposition 14.1.2, ℓ. 1 : Replace 'Let X, Y' by 'Let X, Z'

14 Polynomials of finite codes

- p. 525, ℓ. -6 : Replace '$\tau m = \tau m + \tau' m$' by '$\sigma m = \tau m + \tau' m$'
- p. 526 Example 14.7.3 : The matrices α and β should be

$$\alpha = \left[\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & -1 & 1 & 0 \end{array}\right], \qquad \beta = \left[\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & -1 \\ 0 & -1 & 1 & -2 \\ 0 & -1 & 1 & -2 \end{array}\right].$$

- p. 534 Add to the Notes: The argument of the proof of Theorem 14.7.5 is well-known in group representation theory. The map from V to Ve is called in Green (2007) the *Schur functor*. Theorem 14.7.5 itself has been generalized in Perrin (2013). The statement holds replacing the submonoid generated by a bifix code by a set S such that the minimal automata of S and of its reversal are strongly connected.

Solution of exercises

- p. 543 Solution 3.6.5 : ℓ. 1, replace 'Let $w \in A^*$ be such that' by 'Let $u \in Z'^*$ and $w \in A^*$ be such that'
  ℓ. -1, replace 'We have shown that $w \in U$' by 'We have shown that $u \in U$'.
- p. 544 Solution 3.8.1 : ℓ. 3 of p. 544 : Replace '$v_{n+p} = v_n k^p - \sum_{i=1}^{p} u_{n+i} k^{p-i}$' by '$v_n = k^p - \sum_{i=1}^{n} u_{n-i} k^i$', and insert ℓ. 4, before 'Using' the sentence 'It implies that $v_{n+p} = v_n k^p - \sum_{i=1}^{p} u_{n+i} k^{p-i}$.'
- p. 552 Solution 6.1.2 : replace lines 5–9 by
  'Next, if (ii) holds, consider $x \in H \cap A^*$. Then $x = h_1^{\epsilon_1} h_2^{\epsilon_2} \cdots h_n^{\epsilon_n}$ with $n \geq 0$, $h_i \in X$ and $\epsilon_i = \pm 1$. We may assume that n is chosen minimal. Assume that $\epsilon_i = -1$ for some index i. Since X is bifix, none of the h_i^{-1} can cancel completely with h_{i-1} or with h_{i+1}. Since $x \in A^*$, there exists an index i with $1 \leq i \leq n$ such that $\epsilon_i = -1$ and $h_i^{\epsilon_i}$ cancels with its neighbors, that is $h_{i-1}^{\epsilon_{i-1}} h_i^{\epsilon_i} h_{i+1}^{\epsilon_{i+1}} \in A^*$. Thus, we have $\epsilon_{i-1} = 1$, $\epsilon_{i+1} = 1$ and $h_{i-1} = tu$, $h_i = vu$, $h_{i+1} = vw$ for $t, u, v, w \in A^*$. But then $h_{i-1} h_i^{-1} h_{i+1} = tw$ is in X by (ii). This contradicts the minimality of n. This shows that $\epsilon_i = 1$ for all i and thus $x \in X^*$. Thus (iii) holds.'
- p. 568 Solution 9.3.13 : replace the two last paragraphs by:
  'Let $u = u_s \cdots u_1$ and $v = v_1 \cdots v_t$. We have $|u| \leq (s-1)n(n-1)/2$ and $|v| \leq (t-1)n(n-1)/2$. Thus

  $$|uv| \leq (s+t-2)n(n-1)/2. \tag{2}$$

  Let $z \in A^*$ be such that $q_t \xrightarrow{z} p_s$ with $|z| \leq n-1$. Since $p_s \xrightarrow{u} p_1$ and $q_1 \xrightarrow{v} q_t$, we have $q_1 \xrightarrow{vzu} p_1$. This forces $x_s y_t = 1$ by unambiguity.
  Since $x_s y_t = 1$, we have

  $$s + t \leq \sum_{q \in Q} (x_s)_q + \sum_{q \in Q} (y_t)_q \leq n + 1. \tag{3}$$

  Since the minimal rank of the elements of M is 1, the minimal number of nonzero distinct rows of an element of M is 1. By Exercise 9.3.5, , y_t is a column of an element of the monoid $M = \varphi(A^*)$ with minimal number of nonzero distinct rows. Such an element has the form $m = y_t \ell$ where ℓ is a row vector. Similarly, x_s is a row of an element of M of the form $n = r x_s$ where r is a column vector. Since the minimal rank of the words in $\mathcal{A}$ is 1, we cannot have $\ell r = 0$ which would imply that $0 \in M$. Since $\mathcal{A}$ is unambiguous, this forces $\ell r = 1$ and thus $mn = y_t x_s$. This shows that $y_t x_s \in M$.
  The word $w = vzu$ is such that $y_t x_s \leq \varphi(w)$. Since $y_t x_s \in M$, by Exercise 9.3.12, this implies $y_t x_s = \varphi(w)$. Thus w has rank one and by Equations (2) and (3), $|w| \leq (s+t-2)n(n-1)+n-1 \leq (n^2-n+2)(n-1)/2$.'
- p. 584 Solution 14.1.3 : insert 'strict' before 'right contexts' and 'left contexts'

Appendix: Research problems

- p. 593 ℓ. 8 : Replace the last sentence of the paragraph by 'It is conjectured that for any finite maximal prefix code X there exist $P, T \subset A^*$ such that

$$\underline{X} - 1 = \underline{P}(\underline{A} - 1)\underline{T}$$

where T is the union of $d(X)$ pairwise disjoint maximal prefix sets (see Perrin, Schützenberger (1992)). This is equivalent to say that in Equation (14.7) one has $S = 1$ and the polynomial Q has the form $Q = \sum_{i=1}^{d-1} \underline{U}_i$ where each U_i is a nonempty prefix-closed set.'
- p. 592 Suppress the first sentence (see the complement to page 433).
- p. 593 ℓ. -4 : Replace 'finite set Y' by 'finite subset Y'

References

- p. 596 ℓ. -10 : Replace 'Capoceli' by 'Capocelli'

Index

- p. 613 Add p. 102 for Dyck code.
- p. 616 ℓ. 5 : Replace 'nil-simple semigroup 417' by 'nil-simple semigroup 416'

Additional references

Jorge Almeida (1994). *Finite semigroups and universal algebra*, volume 3 of *Series in Algebra*. World Scientific Publishing Co., Inc., River Edge, NJ, 1994. ISBN 981-02-1895-8. Translated from the 1992 Portuguese original and revised by the author.

Marie-Pierre Béal and Catalin Dima (2015). ℕ-algebraicity of zeta functions of sofic-Dyck shifts. , 2015.

Leonid A. Bokut and Evgeny S. Chibrikov (2006). Lyndon-Shirshov words, Gröbner-Shirshov bases, and free Lie algebras. In *Non-associative algebra and its applications*, volume 246 of *Lect. Notes Pure Appl. Math.*, pages 17–39. Chapman & Hall/CRC, Boca Raton, FL, 2006.

J. A. Green (2007). *Polynomial representations of* GL_n, volume 830 of *Lecture Notes in Mathematics*. Springer, Berlin, augmented edition, 2007. With an appendix on Schensted correspondence and Littelmann paths by K. Erdmann, Green and M. Schocker.

Marshall Hall, Jr. (1976). *The theory of groups*. Chelsea Publishing Co., New York, 1976. Reprinting of the 1968 edition.

Philip Hall (1934). A contribution to the theory of groups of prime-power order. *Proc. London Math. Soc.*, **36**:29–95, 1934.

Gerhard Keller (1991). Circular codes, loop counting, and zeta-functions. *J. Combin. Theory Ser. A*, **56**(1):75–83, 1991.

Michel Lazard (1954). Sur les groupes nilpotents et les anneaux de Lie. *Ann. Sci. Ecole Norm. Sup. (3)*, **71**:101–190, 1954. ISSN 0012-9593.

Yun Liu (2012). Groups and decompositions of codes. *Theor. Comput. Sci.*, **443**:70–81, 2012.

M. Lothaire (1997). *Combinatorics on Words*. Cambridge University Press, second edition, 1997. (First edition 1983).

S. Margolis, M. Sapir, and P. Weil (1998). Irreducibility of certain pseudovarieties. *Comm. Algebra*, **26**(3):779–792, 1998.

Dominique Perrin (2013). Completely reducible sets. *Internat. J. Algebra Comput.*, **23**(4):915–941, 2013.

Christophe Reutenauer (1979). Une topologie du monoïde libre. *Semigroup Forum*, **18**(1):33–49, 1979.

Christophe Reutenauer (1993). *Free Lie Algebras*, volume 7 of *London Mathematical Monographs New Series*. Oxford University Press, 1993.

Robert A. Scholtz (1969). Maximal and variable word-length comma-free codes. *IEEE Trans. Information Theory*, **IT-15**:300–306, 1969.

Imre Simon (1990). Factorization forests of finite height. *Theoret. Comput. Sci.*, **72**(1):65–94, 1990.

Ludwig Staiger (2007). On maximal prefix codes. *Bull. Eur. Assoc. Theor. Comput. Sci. EATCS*, (91):205–207, 2007. ISSN 0252-9742.

Gérard Viennot (1978). *Algèbres de Lie et monoïdes libres*, volume 691 of *Lecture Notes in Mathematics*. Springer-Verlag, 1978.