1

Capitolo C60
parole infinite

Contenuti delle sezioni

8 pagine

C600.01 Questo capitolo è dedicato alle parole infinite sopra un determinato alfabeto A, ossia alle funzioni del genere $\lceil \mathbb{N} \longmapsto \mathsf{A} \rceil$ o del genere $\lceil \mathbb{Z} \longmapsto \mathsf{A} \rceil$.

Si tratta di estensioni tutt'altro che banali delle stringhe finite. In effetti l'insieme di queste funzioni, soprattutto di quelle costruibili, riveste grande importanza per questioni generali attinenti ad algoritmica, computabilità, decidibilità e complessità, a causa dei suoi collegamenti con settori della matematica come semigruppi moltiplicativi di matrici, teoria dei gruppi e probabilità e per rilevanti applicazioni negli ambiti della fisica, della biologia e dell'informatica applicata.

Nelle prime pagine si introducono le definizioni e le costruzioni basilari per lo studio generale delle parole infinite.

Successivamente si procede allo studio di specifiche parole infinite, come la coppia delle parole di Morse-Thue, la parola di Fibonacci e la parola tribonacci. Si trattano poi le sottoclassi delle parole periodiche e delle prole bilanciate.

Nelle pagine finali si esaminano con una certa ampiezza le parole sturmiane.

C60 a. parole infinite [1]

C60a.01 Per vari problemi risulta utile studiare sequenze infinite di oggetti elementari costituenti un insieme finito. Per avere risultati generali questi vari oggetti vanno rappresentati dai caratteri facenti parte di un alfabeto (finito) A.

In questo capitolo chiamiamo **parola infinita** [a destra] sull'alfabeto A una funzione w del genere $\lceil \mathbb{N} \longmapsto \mathsf{A} \rceil$ che sia costruibile, cioè per la quale esista un procedimento che per ogni $j \in \mathbb{N}$ consenta di individuare il carattere di A costituente la sua componente j-esima che denotiamo con c_j.

A una tale w daremo anche la forma della stringa numerabile e la forma della successione scrivendo uguaglianze come

$$\mathsf{w} \;=\; \left\downarrow \begin{matrix} 0 & 1 & 2 & \ldots & n & \ldots \\ c_0 & c_1 & c_2 & \ldots & c_n & \ldots \end{matrix} \right\downarrow \;=\; c_0 \,{\scriptstyle\mathsf{I}}\, c_1 \,{\scriptstyle\mathsf{I}}\, c_2 \,{\scriptstyle\mathsf{I}}\, \ldots \,{\scriptstyle\mathsf{I}}\, c_n \ldots \;=\; \langle c_0, c_1, c_2, ..., c_n, ... \rangle \;.$$

Per i suoi caratteri componenti possiamo usare le notazioni

$$\forall j \in \mathbb{N} \;\; \vdots \;\; [\mathsf{w}]_j \;:=\; \boldsymbol{Prj}_j(\mathsf{w}) \;:=\; \mathsf{w}(j) = c_j \;.$$

L'insieme $\mathsf{A}^{\mathbb{N}}$ della parole infinite sull'alfabeto A viene denotato spesso con A^ω .

Denoteremo invece con A^∞ l'insieme delle stringhe su A finite o infinite , cioè poniamo

$$\mathsf{A}^\infty \;:=\; \mathsf{A}^* \,\dot\cup\, \mathsf{A}^\omega \;.$$

Nel seguito continueremo ad usare il termine stringhe per le funzioni finite con valori in un alfabeto e spesso semplificheremo il termine "parola infinita" con **parola-N**.

C60a.02 Per le funzioni del genere $\lceil \mathbb{P} \longmapsto \mathsf{A} \rceil$ o di un genere della forma $\lceil \{z, z{+}1, z{+}2, ...\} \longmapsto \mathsf{A} \rceil$ per qualche $z \in \mathbb{Z}$ si possono svolgere argomentazioni e ottenere asserzioni derivabili da quelle esposti di seguito attraverso modifiche di notazioni prive di difficoltà.

Occorre segnalare che si affrontano anche molti problemi che richiedono di conoscere proprietà delle funzioni del genere $\lceil \mathbb{Z} \longmapsto \mathsf{A} \rceil$; queste entità sono dette **parole infinite bilatere** o **parole-Z** su A e il loro insieme, evidentemente, si può denotare con $\mathsf{A}^{\mathbb{Z}}$.

Segnaliamo anche che varie questioni enumerative si affrontano esaminando sequenze finite e infinite composte da oggetti elementari estratti da insiemi numerabili che vengono detti **alfabeti infiniti**.
Viene dunque sviluppata anche una teoria delle stringhe finite e infinite su alfabeti infiniti [*Goulden, Jackson (1983)*].

Introduciamo anche il termine **funzione-LtN** come abbreviazione di funzione di un genere $\lceil \mathsf{L} \longmapsto \mathbb{N} \rceil$ riguardante un qualche linguaggio L su un qualche alfabeto A.
La lunghezza delle stringhe è una funzione di questo genere.

Abbreviamo inoltre il termine funzione di un genere $\lceil \mathsf{L} \longmapsto \mathsf{M} \rceil$ per qualche coppia di linguaggi L e M su uno o due alfabeti con **funzione-LtL**.
Con abbreviazioni simili si possono individuare funzioni tra estensioni booleane di linguaggi e insiemi numerici: ad esempio "funzioni-LBtNB" abbrevia "funzioni da insiemi di linguaggi in insiemi di naturali", mentre "funzioni-LBtRB" sta per "funzioni da insiemi di linguaggi in insiemi di numeri reali".

C60a.03 Lo studio delle parole infinite si rivolge primariamente a quelle su un alfabeto di due caratteri (le parole infinite su una sola lettera sono essenzialmente sottoinsiemi di $\mathbb{N}$ o di $\mathbb{Z}$).

In molte applicazioni all'informatica si considerano parole binarie, parole sull'alfabeto $\{0, 1\}$ o anche parole sopra l'insieme dei due valori di verità vero e falso, $\{\mathbf{t}, \mathbf{f}\rangle$.

Qui di seguito ci occuperemo prevalentemente delle parole sugli alfabeti $\{a,b\}$ e $\{0,1\}$.

Molte parole infinite si individuano a partire da successioni di stringhe $\mathbf{u} = \langle u_0, u_1, u_2, ..., u_n, ...\rangle$ aventi lunghezze crescenti e tali che $\forall i \in \mathbb{N} :\ u_i \prec_{\mathsf{p}} u_{i+1}$.

Si dice **limite della successione di parole u** la parola infinita $\mathbf{p}$ la cui componente $\mathbf{p}(s)$ coincide con la componente $p_i(s)$ di tutte le stringhe p_i aventi lunghezza superiore ad s.
Nel seguito presenteremo alcune di tali parole le quali presentano proprietà di notevole eleganza.

C60a.04 Definiamo il **morfismo di monoidi** $\tau \in \lceil \mathsf{A}^* \longmapsto \mathsf{A}^* \rceil$ chiedendo

$$\tau(a) := a\,b \qquad\qquad \tau(b) := b\,a\ .$$

Applicando reiteratamente τ alle lettere a e b si ottengono le seguenti due sequenze di stringhe:

$$\begin{array}{ll} \tau(a) = ab & \tau(b) = ba \\ \tau^2(a) = abba & \tau^2(b) = baab \\ \tau^3(a) = abbabaab & \tau^3(b) = baababba \\ \tau^4(a) = abbabaabbaababba & \tau^4(b) = baababbaabbabaab \end{array}$$

Le proprietà di queste stringhe si esprimono facilmente servendosi dell'isomorfismo $\sigma \in \lceil \mathsf{A}^* \triangleleft\!\!\longrightarrow\!\!\triangleright \mathsf{A}^* \rceil$ definito come estensione per giustapposizione dello scambio

$$\sigma(a) = b, \qquad \sigma(b) = a\ ,$$

che evidentemente è un'involuzione entro A^*.

C60a.05 Prop. $\forall s = 0,1,2,...$:
$\sigma(\tau^s(a)) = \tau^s(b), \qquad \sigma(\tau^s(b)) = \tau^s(a),$
$\tau^{s+1}(a) = \tau^s(a)\tau^s(b), \qquad \tau^{s+1}(b) = \tau^s(b)\tau^s(a), \qquad |\tau^s(a)| = |\tau^s(b)| = 2^s,$
$(\tau^{2s}(a))^{\leftarrow} = \tau^{2s}(a), \qquad (\tau^{2s}(b))^{\leftarrow} = \tau^{2s}(b), \qquad \sigma(\tau^{2s+1}(a)) = \tau^{2s+1}(b).$

Dim.: Le uguaglianze si provano senza difficoltà procedendo per induzione ▮

Le due precedenti sequenze, dato che ciascuna delle sue stringhe è prefisso delle successiva, definiscono due parole infinite che denotiamo con $\mathbf{t}$ e $\sigma(\mathbf{t})$; $\mathbf{t}$ è detta **parola infinita di Thue-Morse**, mentre $\sigma(\mathbf{t})$ è chiamata **coniugata della t**. Si osservi che da $\tau^s(a)$ e $\tau^s(b)$ si possono conoscere le prime 2^s componenti, risp., di $\mathbf{t}$ e $\sigma(\mathbf{t})$.

Nella parola di Thue-Morse si trovano come fattori stringhe al quadrato: dimostriamo che in essa non si trovancome fattori né coppie sovrapposte, né stringhe al cubo.

C60a.06 Prop. Se $\mathsf{X} = \{ab, ba\}$ ed $x \in \mathsf{X}^*$, allora $axa, bxb \notin \mathsf{X}^*$.

Dim.: Dimostriamo le relazioni di non appartenenza per induzione su $x^{\vdash\!\dashv}$. Esse sono ovvie per $x = \mu$. Supponiamole vere per $x^{\vdash\!\dashv} = 2r$ e consideriamo $w = axa = w_1..w_{r+1}$ con $w_i \in \mathsf{X}$; deve essere $w_1 = ab$ e $w_{r+1} = ba$ e $v = w_2...w_r \in \mathsf{X}^*$. 5

C60 b. parole di Fibonacci

C60b.01 Consideriamo $\mathsf{A} := \{a, b\}$ e il morfismo $\phi_F \in \lceil \mathsf{A}^* \longmapsto \mathsf{A}^* \rceil$ definito da

$$\phi_F(a) := ab \quad , \qquad \phi_F(b) = a$$

ed introduciamo la successione delle **parole finite di Fibonacci**

$$\mathsf{Ffw} := \langle i \in \mathbb{N} :| \ \mathtt{f}_i \rangle \quad \text{con}$$

$$\mathtt{f}_0 := b, \quad \mathtt{f}_1 := a, \quad \forall s = 0, 1, 2, ... \ \vdots \ \ \mathtt{f}_{s+1} := \phi_F(f_s).$$

Più dettagliatamente si ha:

$$\mathtt{f}_0 = b, \ \ \mathtt{f}_1 = a, \ \ \mathtt{f}_2 = ab, \ \ \mathtt{f}_3 = aba, \ \ \mathtt{f}_4 = abaab, \ \ \mathtt{f}_5 = abaababa,$$

$$\mathtt{f}_6 = abaababaabaab, \ \ \mathtt{f}_7 = abaababaabaababaababa,$$

$$\mathtt{f}_8 = abaababaabaababaababaabaababaabaab, ... \ .$$

C60b.02 Prop.

$$\forall s = 2, 3, 4, ... \ \vdots \ \ \mathtt{f}_s = \mathtt{f}_{s-1} \cdot \mathtt{f}_{s-2}.$$

Dim.: Verificato che $\mathtt{f}_2 = \mathtt{f}_1 \cdot \mathtt{f}_0$, si procede per induzione supponendo sia $\mathtt{f}_s = \mathtt{f}_{s-1} \cdot \mathtt{f}_{s-2}$; applicando il morfismo ϕ_F si ha $\phi_F(\mathtt{f}_s) = \phi_F(\mathtt{f}_{s-1}) \cdot \phi_F(\mathtt{f}_{s-2})$, cioè $\mathtt{f}_{s+1} = \mathtt{f}_s \cdot \mathtt{f}_{s-1}$ ▮

Quindi ciascuna parola $\mathtt{f}_s$ è prefisso della successiva $\mathtt{f}_{s+1}$ e si può definire la parola infinita limite anche per questa sequenza. Questa è detta **parola infinita di Fibonacci** e verrà qui denotata con Fiw.

Il nome delle parole ora introdotte si collega alla successione dei **numeri di Fibonacci** definita da:

$$\mathsf{FibN} := \langle i \in \mathbb{N} :| \ \mathtt{F}_i \rangle \quad \text{con}$$

$$\mathtt{F}_0 = \mathtt{F}_1 = 1; \quad \forall s = 2, 3, 4, ... \ \vdots \ \ \mathtt{F}_s = \mathtt{F}_{s-1} + \mathtt{F}_{s-2},$$

C60b.03 Prop. Le lunghezze delle parole finite di Fibonacci sono i numeri di Fibonacci: $\mathtt{F}_s := \mathtt{f}_s{}^{\vdash\dashv}$.

Dim.: Osservato che $\mathtt{f}_0{}^{\vdash\dashv} = \mathtt{f}_1{}^{\vdash\dashv} = 1$, dalla $\mathtt{f}_s = \mathtt{f}_{s-1} \cdot \mathtt{f}_{s-2}$ si ricava la relazione di ricorrenza $\mathtt{f}_s{}^{\vdash\dashv} = \mathtt{f}_{s-1}{}^{\vdash\dashv} + \mathtt{f}_{s-2}{}^{\vdash\dashv}$, ossia la $\mathtt{F}_s = \mathtt{F}_{s-1} + \mathtt{F}_{s-2}$; quindi l'asserto risulta dimostrato per induzione ▮

Molte proprietà delle parole di Fibonacci si dimostrano facilmente procedendo per induzione. Per esempio:

C60b.04 Prop. Per $s = 2, 3, 4, ...$ si ha $\mathtt{f}_s{}^{\mathcal{P}rk} = \langle \mathtt{F}_{s-1}, \mathtt{F}_{s-2} \rangle$.

Verificata l'uguaglianza per $\mathtt{f}_2$, supponiamola vera per $\mathtt{f}_s$ e consideriamo $\mathtt{f}_{s+1}$:
$\mathtt{f}_{s+1}{}^{\mathcal{P}rk} = (\mathtt{f}_s \mathtt{f}_{s-1})^{\mathcal{P}rk} = (\mathtt{f}_s)^{\mathcal{P}rk} \mathbf{+} \mathtt{f}(\mathtt{f}_{s+1})^{\mathcal{P}rk} = \langle \mathtt{F}_{s-1}, \mathtt{F}_{s-2} \rangle \mathbf{+} \langle \mathtt{F}_{s-2}, \mathtt{F}_{s-3} \rangle = \langle \mathtt{F}_s, \mathtt{F}_{s-1} \rangle$ e l'asserto risulta dimostrato per induzione ▮

Per $s = 2, 3, ...$ poniamo in evidenza le ultime due lettere delle parole di Fibonacci scrivendo $\mathtt{f}_s = h_s c_s d_s$ con $c_s, d_s \in \{a, b\}$ e definendo $h_s := \mathtt{f}_s /\!\!/ 2$.

C60b.05 (1) Prop.: Sia $r = 1, 2, ...$; ogni parola di Fibonacci relativa ad indice pari $\mathtt{f}_{2r}$ termina con $c_{2r} d_{2r} = ab$ e ogni parola relativa a indice dispari $\mathtt{f}_{2r+1}$ termina con $c_{2r+1} d_{2r+1} = ba$ ▮

(2) Prop.: Nelle parole di Fibonacci ogni occorrenza di b o è finale di parola di indice pari o è preceduta e seguita da a; ogni occorrenza di a^2 è preceduta e seguita da b ▮

(3) Prop.: a^3 e b^2 non sono infissi di alcuna parola di Fibonacci finita o infinita ▮

Per $s = 2, 3, ...$ scriviamo $g_s := f_{s-2} f_{s-1}$ e confrontiamo le prime componenti di questa successione con le corrispondenti h_s:

$$g_2 = ba \quad g_3 = aab \quad g_4 = ababa \quad g_5 = abaabaab \quad g_6 = abaababaababa$$
$$h_2 = \mu \quad h_3 = a \quad h_4 = aba \quad h_5 = abaaba \quad h_6 = abaababaaba.$$

C60b.06 Prop. $\forall s = 2, 3, ...$:
(a) $f_s = f_{s-1} f_{s-2} = f_{s-2} f_{s-3} f_{s-2} = f_{s-2} g_{s-1} = h_s c_s d_s$;
(b) $g_s = f_{s-2} f_{s-1} = f_{s-2} f_{s-3} g_{s-2} = f_{s-1} g_{s-2} = h_s d_s c_s$;
(c) il massimo prefisso comune ad f_s e g_s è h_s ;
(d) $h_{s+2} = f_{s+1} h_s = f_s h_{s+1} = h_s {f_{s+1}}^{\leftarrow} = h_{s+1} {f_s}^{\leftarrow}$;
(e) $h_{s+3} = {f_{s+1}}^2 h_s$;
(f) $h_{s+3} = h_{s+1} d_s c_s h_s c_s d_s h_{s+1}$;
(g) ${h_s}^{\leftarrow} = h_s$;
(h) $\forall s = 2, 3, ..., \; t = 1, 2, ...$: h_s è prefisso e suffisso di h_{s+t} ;
(i) $h_{2r} = f_{2r-1} f_{2r-3} ... f_3, \quad h_{2r+1} = f_{2r} f_{2r-2} ... f_4 f_1$.

Dim.: (a) applicando due volte b05(2) si hanno le prime due uguaglianze; la definizione di g_s conduce alla terza; la definizione di h_s alla quarta.

(b) Dopo la definizione di g_s, si ha una uguaglianza derivante dalla $f_s = f_{s-2} g_{s-1}$ in (a); la terza si ottiene dab05(2); la quarta si ottiene per induzione servendosi della $g_s = f_{s-1} g_{s-2}$.

(c) Si ricava dal confronto di (a) e (b).

(d) $h_{s+2} = \lfloor_{\mathsf{N}}$ definizione $\rfloor_{\mathsf{N}} = f_{s+2} /\!\!/ 2 = \lfloor_{\mathsf{N}}$ b05(2) $\rfloor_{\mathsf{N}} = f_{s+1} f_s /\!\!/ 2 = f_{s+1} h_s$;
per la seconda uguaglianza $f_{s+1} f_s /\!\!/ 2 = f_s f_{s-1} f_s /\!\!/ 2 = f_s g_{s+1} /\!\!/ 2$;
terza e quarta uguaglianza si ricavano dalle precedenti e da (g).

(e) $h_{s+3} = \lfloor_{\mathsf{N}}$ (d) $\rfloor_{\mathsf{N}} = f_{s+1} h_{s+2} = f_{s+1} f_{s+1} h_s$.

(f) $h_{s+3} = \lfloor_{\mathsf{N}}$ (d) $\rfloor_{\mathsf{N}} = f_{s+1} h_{s+2} = \lfloor_{\mathsf{N}}$ definizione e (d) $\rfloor_{\mathsf{N}} = h_{s+1} c_{s+1} d_{s+1} f_s h_{s+1} = h_{s+1} d_s c_s h_s c_s d_s h_{s+1}$.

(g) Si verifica per i primi h_s e si ricava per induzione da (f).

(h) $g_s =$ [definizione] $= f_{s-1} g_{s-2} =$ [(a)] $= h_{s-1} c_{s-1} d_{s-1} g_{s-2}$.

Di conseguenza $h_s = g_s /\!\!/ 2 = h_{s-1} c_{s-1} d_{s-1} g_{s-2} /\!\!/ 2 = h_{s-1} c_{s-1} d_{s-1} h_{s-2}$; da qui $h_{s-1} \prec_{\mathsf{P}} h_s$ e quindi $h_s \prec_{\mathsf{P}} h_{s+t}$; dunque $h_s \prec_{\mathsf{S}} h_{s+t}$ scende direttamente da (d) .

(i) $f_s = f_{s-1} f_{s-2} = f_{s-1} f_{s-3} f_{s-4} = f_{s-1} f_{s-3} f_{s-5} f_{s-6}$; quindi:
$f_{2r} = f_{2r-1} f_{2r-3} ... f_3 f_2$, da cui la prima uguaglianza;
$f_{2r+1} = f_{2r} f_{2r-2} ... f_4 f_3$, da cui la seconda ▮

C60b.07 Ricordiamo che nell'ambito della teoria dei codici [C65, C66] per **codifica binaria** si intende un isomorfismo del genere $\lceil \mathsf{A}^* \lhd\!\!\longrightarrow \{0, 1\}^* \rceil$ il quale consente di esprimere mediante cifre binarie le più svariate informazioni per poterle trattare con gli odierni strumenti dell'informatica e delle telecomunicazioni senza perdere la possibilità di ricostruire il loro significato originario.

In particolare si hanno le codifiche binarie dei caratteri utilizzati correntemente con il computer seguendo il cosiddetto codice ASCII, e quelle riguardanti decine di migliaia di segni degli alfabeti delle lingue di qualche importanze secondo il codice `Unicode (we)`.

Eserc. Per la parola di Thue-Morse $\mathbf{t} = \mathbf{t}(0)\mathbf{t}(1)\mathbf{t}(2)...$, dimostrare che per $s = 0, 1, 2, ...$ abbiamo $\tau(\mathbf{t}(s)) = \mathbf{t}(2s)\mathbf{t}(2s+1)$ e quindi $\tau(\mathbf{t}) = \mathbf{t}$. Dedurre poi che per ogni s si ha $\mathbf{t}(s) = a$ sse nella scrittura binaria di s si ha un numero pari di cifre 1, mentre $\mathbf{t}(s) = b$ nel caso contrario.

C60 c. parole bilanciate

C60c.01 Ricordiamo che, se w denota una parola sopra un qualche alfabeto A, con $\mathsf{Ftr}(\mathsf{w})$ denotiamo il linguaggio costituito dai fattori, ossia dagli infissi, di tale parola.
Inoltre identifichiamo Ftr con la sua estensione booleana, e quindi se L è un linguaggio usiamo la notazione $\mathsf{Ftr}(\mathsf{L}$ per il linguaggio dei fattori delle parole in L.

Per ogni $n \in \mathbb{N}$ introduciamo anche le notazioni $\mathsf{Ftr}_n(\mathsf{w}) := \mathsf{Ftr}(\mathsf{w}) \cap \mathsf{A}^n$ e $\mathsf{Ftr}_n(\mathsf{L}) := \mathsf{Ftr}(\mathsf{L}) \cap \mathsf{A}^n$.

Se L è un linguaggio denotiamo con $\boldsymbol{g}_{\mathsf{L}}$ la funzione che ad ogni $n \in \mathbb{N}$ associa il numero delle stringhe di L di lunghezza n; si tratta quindi di una funzione-NtN.

Questa funzione viene detta **funzione di enumerazione del linguaggio** L o **funzione della complessità-u del linguaggio** L. Il termine complessità-u va considerato abbreviazione di "complessità delle sottoparole".

Invece della notazione $\boldsymbol{g}_{\mathsf{Ftr}(\mathsf{w})}$ useremo anche la sua abbreviazione $\boldsymbol{g}_{\mathsf{w}}$; quindi possiamo scrivere

$$\boldsymbol{g}_{\mathsf{w}}(n) := |\mathsf{Ftr}(\mathsf{w}) \cap \mathsf{A}^n| .$$

Ogni fattore $u \in \mathsf{Ftr}(\mathsf{w})$ si dice prolungabile a sinistra sse esiste una lettera $a_j \in \mathsf{A}$ tale che $a_j\, u \in \mathsf{Ftr}(\mathsf{w})$; chiaramente ogni $u \in \mathsf{Ftr}(\mathsf{w}) \setminus \mathsf{Pfx}(\mathsf{w})$ è prolungabile a sinistra.

Dualmente-LR ogni $u \in \mathsf{Ftr}(\mathsf{w})$ si dice prolungabile a destra sse esiste almeno una lettera $a_j \in \mathsf{A}$ tale che $u\, a_j \in \mathsf{Ftr}(\mathsf{w})$. Questo accade ad ogni fattore u, dato che ogni $u \in \mathsf{Ftr}(\mathsf{w})$ è prefisso immediato di almeno un altro fattore. Di conseguenza ogni $\boldsymbol{g}_{\mathsf{w}}(n)$ è una funzione-LtN nondecrescente e quindi

$$\forall n \in \mathbb{N} : \ \boldsymbol{g}_{\mathsf{w}}(n) \leq \boldsymbol{g}_{\mathsf{w}}(n+1) .$$

C60c.02 Un fattore u della w si dice **fattore conservativo della parola** sse è prolungabile a destra con una sola lettera; si dice invece **fattore espansivo della parola** sse si può prolungare con 2 o più lettere.

Denotiamo con $\boldsymbol{e}_{\mathsf{w}}(u)$ il numero delle lettere $a_j \in \mathsf{A}$ per le quali $u\, a_j \in \mathsf{Ftr}(\mathsf{w})$. Per ogni $\mathsf{w} \in \mathsf{A}^\omega$ $\boldsymbol{e}_{\mathsf{w}}$ è una funzione-LtL a valori positivi. Evidentemente un fattore è conservativo sse $\boldsymbol{e}_{\mathsf{w}}(u) = 1$.

Introduciamo per ogni w la funzione-NtN

$$\boldsymbol{E}_{\mathsf{w}} := \lceil n \in \mathbb{N} \mapsto \sum_{u \in \mathsf{Ftr}_n(\mathsf{w})} (\boldsymbol{e}_{\mathsf{w}}(u) - 1) \rceil .$$

Chiaramente:

$$\forall n \in \mathbb{N} : \ \boldsymbol{g}_{\mathsf{w}}(n+1) = \boldsymbol{g}_{\mathsf{w}}(n) + \boldsymbol{E}_{\mathsf{w}}(n) .$$

Nel caso in cui l'alfabeto è costituito da due lettere $\boldsymbol{E}_{\mathsf{w}}(n)$ fornisce il numero dei fattori di w che sono espansivi e di lunghezza n.

Una parola-N su A si dice **parola definitivamente periodica** sse esistono una stringa $u \in \mathsf{A}^*$ e una stringa $v \in \mathsf{A}^+$ tali che si possa scrivere $\mathsf{w} = u\, v\, v \ldots v \ldots$.

Spesso il secondo membro della precedente uguaglianza viene denotato con $u\, v^\omega$.

C60c.03 Morse e Hedlund hanno dimostrata la seguente caratterizzazione delle parole definitivamente periodiche.

(1) Prop.: Consideriamo la parola infinita w e sia $\mathsf{A} := \text{minalf}(\mathsf{w}) = k$. Gli enunciati che seguono sono equivalenti.

(a) w è definitivament eperiodica ;

(b) la funzione di complessità $\boldsymbol{g}_{\mathbf{w}}(s)$ è definitivamente costante ;

(c) esiste $s \in \mathbb{P}$ tale che $\boldsymbol{g}_{\mathbf{w}}(s) \leq s + k - 2$;

(d) esiste $s \in \mathbb{P}$ tale che $\boldsymbol{g}_{\mathbf{w}}(s) = \boldsymbol{g}_{\mathbf{w}}(s + 1)$ ▮

C60c.04 Una parola-N **w** si dice **parola ricorrente** sse ogni suo fattore u si trova nella **w** in un numero infinito di occorrenze.

Una parola infinita **w** si dice **parola uniformemente ricorrente** sse esiste una funzione-LtN $\mathcal{K} \in \lceil \mathsf{Ftr}(\mathbf{w}) \longmapsto \mathbb{N} \rceil$ tale che per ogni $u, v \in \mathsf{Ftr}(\mathbf{w})$ con $|v| \geq \mathcal{K}(u)$ accade che $u \in \mathsf{Ftr}(v)$.

C60c.05 Consideriamo un alfabeto di due lettere $\mathsf{A} = \{a, b\}$.

Per ogni coppia di parole u e v su A aventi la stessa lunghezza si dice **sbilanciamento della coppia di parole** il numero naturale

$$\Big| |u|_a - |v|_a \Big| \;=\; \Big| |u|_b - |v|_b \Big| \;.$$

C60 d. parole episturmiane

C60d.01 Anche delle parole sturmiame si cercano generalizzazioni, in particolare generalizzazioni costituite da parole su tre o più lettere chiamate **parole episturmiane**.

Un esempio è costituito dalle cosiddetta **parola tribonacci** che si possono considerare anche generalizzazioni della parola di Fibonacci.

Anche questa parola-N che denotiamo con **t** è definita come il limite di una successione di stringhe definite mediante una relazione di ricorrenza:

$$t_0 \ := \ 0 \ , \ t_1 \ := \ 01 \ , \ t_2 \ := \ 0102 \ , \ t_{n+3} \ := \ t_{n+2}\, t_{n+1}\, t_n \ .$$

Per la parola tribonacci possiamo quindi scrivere

$$\mathbf{t} \ = \ 01020100102010201\dots$$

Essa si può equivalentemente definire come il punto fisso del morfismo $\left\downarrow\begin{array}{ccc} 0 & 1 & 2 \\ 01 & 02 & 0 \end{array}\right\downarrow$ e quindi scrivere:

$$\mathbf{t} \ := \ (0102)_{\urcorner} \left\downarrow\begin{array}{ccc} 0 & 1 & 2 \\ 01 & 02 & 0 \end{array}\right\downarrow^{\oplus} \ .$$

C60d.02 Segnaliamo altre parole infinite generate iterando morfismi.

Il morfismo $\left\downarrow\begin{array}{ccc} a & b & c \\ ab & bc & c \end{array}\right\downarrow$ porta alla parola

$$\mathbf{t} \ := \ (abbc)_{\urcorner} \left\downarrow\begin{array}{ccc} a & b & c \\ ab & bc & c \end{array}\right\downarrow^{\oplus} \ = \ abbcbc^2bc^3\dots bc^n \dots \ .$$

Il morfismo $\left\downarrow\begin{array}{ccc} a & b & c \\ abc & bb & ccc \end{array}\right\downarrow$ porta alla parola

$$a_{\urcorner} \left\downarrow\begin{array}{ccc} a & b & c \\ ab & bc & c \end{array}\right\downarrow^{\oplus} \ = \ abcb^2c^3b^4c^9\dots b^{2^n}\, c^{3^n} \dots \ .$$

Il morfismo $\left\downarrow\begin{array}{cc} a & b \\ abab & bb \end{array}\right\downarrow$ porta alla parola

$$(ab)_{\urcorner} \left\downarrow\begin{array}{cc} a & b \\ abab & bb \end{array}\right\downarrow^{\oplus} \ = \ abab^3\, abab^7\, ababbbabab^{15} \dots \ .$$