Attractors of sequences coding β-integers

Ľubomíra Dvořáková, Martina Moravcová

Department of Mathematics, FNSPE Czech Technical University in Prague

June, 9, 2026

Program

Program

String attractor

Let $w = w_0 w_1 \cdots \underline{w_i w_{i+1} \cdots w_j} \cdots w_{n-1}$. Then $\{i, i+1, \ldots, j\}$ is an *occurrence* of the factor $z = w_i w_{i+1} \cdots w_j$ in w.

Definition (Kempa & Prezza, 2018)

Let $w = w_0 w_1 \cdots w_{n-1}$. Then $\Gamma \subset \{0, 1, \ldots, n-1\}$ is a *(string) attractor* of w if each factor of w has an occurrence containing an element of Γ.

String attractor

Let $w = w_0 w_1 \cdots \underline{w_i w_{i+1} \cdots w_j} \cdots w_{n-1}$. Then $\{i, i+1, \ldots, j\}$ is an *occurrence* of the factor $z = w_i w_{i+1} \cdots w_j$ in w.

Definition (Kempa & Prezza, 2018)

Let $w = w_0 w_1 \cdots w_{n-1}$. Then $\Gamma \subset \{0, 1, \ldots, n-1\}$ is a *(string) attractor* of w if each factor of w has an occurrence containing an element of Γ.

Example

$$w = \mathrm{ATGCAATGCATG}$$

$$\Gamma = \{2,3,4,8,10\} \to w = \mathrm{AT\underline{GCA}ATG\underline{C}A\underline{T}G}$$

$$\Gamma^* = \{2,3,4,10\} \to w = \mathrm{AT\underline{GCA}ATGCA\underline{T}G}$$

Application in data compression (e.g. bioinformatics)

String attractor

Let $w = w_0 w_1 \cdots \underline{w_i w_{i+1} \cdots w_j} \cdots w_{n-1}$. Then $\{i, i+1, \ldots, j\}$ is an *occurrence* of the factor $z = \overline{w_i w_{i+1} \cdots w_j}$ in w.

Definition (Kempa & Prezza, 2018)

Let $w = w_0 w_1 \cdots w_{n-1}$. Then $\Gamma \subset \{0, 1, \ldots, n-1\}$ is a *(string) attractor* of w if each factor of w has an occurrence containing an element of Γ.

Example

$$w = \mathrm{ATGCAATGCATG}$$

$$\begin{aligned} \Gamma &= \{2,3,4,8,10\} & \to w &= \mathrm{AT\underline{GCA}ATG\underline{C}A\underline{T}G} \\ \Gamma^* &= \{2,3,4,10\} & \to w &= \mathrm{AT\underline{GCA}ATGCA\underline{T}G} \end{aligned}$$

Application in data compression (e.g. bioinformatics)

Program

Known results

In general, to find a minimal attractor is an NP-complete problem.

In CoW, attractors of minimum size have been determined for particular prefixes of the following sequences:

- standard Sturmian by Mantaci, Restivo, Romana, Rosone, Sciortino, 2021
- Thue-Morse by Kutsukake et al., 2020
- CS Rote by Dvořáková & Hendrychová, 2024

Known results

In general, to find a minimal attractor is an NP-complete problem.
In CoW, attractors of minimum size have been determined for particular prefixes of the following sequences:

- standard Sturmian by Mantaci, Restivo, Romana, Rosone, Sciortino, 2021
- Thue-Morse by Kutsukake et al., 2020
- CS Rote by Dvořáková & Hendrychová, 2024

Known results

In CoW, attractors of minimum size have been determined for prefixes of the following sequences:

- standard Sturmian by Restivo et al., 2022
- Tribonacci by Schaeffer & Shallit, 2021
- Thue-Morse by Schaeffer & Shallit, 2021
- period-doubling by Schaeffer & Shallit, 2021
- powers of two by Schaeffer & Shallit, 2021
- **simple Parry** by Gheeraert, Romana, Stipulanti, 2024;
  by Dvořáková & Moravcová, 2025

Known results

In CoW, attractors of minimum size have been determined for <u>factors</u> of the following sequences:

- Thue-Morse by Dolce, 2023
- episturmian by Dvořáková, 2024

Known results

- Schaeffer & Shallit, 2021: study of attractors in linearly recurrent and in automatic sequences
- Béaur, Gheeraert, Hellouin de Menibus, 2024: string attractors and bi-infinite words
- Cassaigne, Gheeraert, Restivo, Romana, Sciortino, Stipulanti, 2025: combinatorial properties of attractors (relation to factor complexity, recurrence function, etc.), new string attractor-based complexities, study of attractors in fixed points of morphisms
- Béal, Crochemore, Romana, 2025: checking and producing word attractors

Program

Techniques used for determining minimal attractors

- $t_{n+1} = t_n \overline{t_n}$ (Thue-Morse sequence)
- WALNUT (automatic sequences)
- palindromic closure (episturmian sequences)
- pseudopalindromic closure (CS Rote sequences)

Example

Let w be a word, then its *palindromic closure* w^R is the shortest palindrome with the prefix w.
$(000)^R = 000, \quad (012)^R = 01210, \quad (0121)^R = 01210.$
Let w be a binary word, then its *antipalindromic closure* w^E is the shortest antipalindrome with the prefix w.
$(000)^E = 000111, \quad (0101)^E = 0101, \quad (0010)^E = 001011.$

- morphism (d-bonacci sequence, Parry sequences)

Motivation

F. Gheeraert, G. Romana, M. Stipulanti, *String attractors of some simple-Parry automatic sequences*, TOCS **68** (2024), 1601–1621

- found attractors of prefixes of simple Parry sequences
- **conjectured these attractors are in general not minimal**
- **opened a question on attractors of prefixes of non-simple Parry sequences**

L. Dvořáková, M. Moravcová, *Attractors of sequences coding β-integers*, 2025, `https://arxiv.org/abs/2511.00650`

- **confirmation of the conjecture – finding minimal attractors of prefixes of simple Parry sequences**
- **partial answer for binary non-simple Parry sequences – finding minimal attractors of particular prefixes**

Motivation

F. Gheeraert, G. Romana, M. Stipulanti, *String attractors of some simple-Parry automatic sequences*, TOCS **68** (2024), 1601–1621

- found attractors of prefixes of simple Parry sequences
- **conjectured these attractors are in general not minimal**
- **opened a question on attractors of prefixes of non-simple Parry sequences**

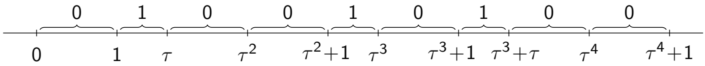
L. Dvořáková, M. Moravcová, *Attractors of sequences coding β-integers*, 2025, `https://arxiv.org/abs/2511.00650`

- **confirmation of the conjecture – finding minimal attractors of prefixes of simple Parry sequences**
- **partial answer for binary non-simple Parry sequences – finding minimal attractors of particular prefixes**

Parry sequence and Rényi numeration systems

- given a base $\beta \in \mathbb{R}$, $\beta > 1$, and $x \in \mathbb{R}$, $x \geq 0$
- $x = \sum_{i=-\infty}^{k} x_i \beta^i \Rightarrow x_k \cdots x_1 x_0 \bullet x_{-1} x_{-2} \cdots$ is a *β-representation* of x
- *β-expansion* of x = β-representation of x obtained by the greedy algorithm, notation $\langle x \rangle_\beta$
- $\langle x \rangle_\beta = x_k x_{k-1} \cdots x_0 \bullet \Rightarrow x$ is a non-negative *β-integer*
- $\mathbb{Z}_\beta^+$ = set of non-negative β-integers
- finitely many distances between neighbors in $\mathbb{Z}_\beta^+ \Rightarrow \beta$ is a *Parry number*
- encoding the same distances with the same letters $\Rightarrow$ a *Parry sequence* $\mathbf{u}_\beta$ is obtained

$$\mathbb{Z}_\tau^+ = \{0, 1, \tau, \tau^2, \tau^2+1, \tau^3, \tau^3+1, \tau^3+\tau, \tau^4, \tau^4+1, \dots\}$$



Simple Parry sequence – definition

- Fabre, 1995: Parry sequences are fixed points of morphisms

Definition (Simple Parry sequence)

A *simple Parry sequence* $\mathbf{u} = \mathbf{u}_\beta$ is the fixed point of the morphism

$$\begin{aligned} \varphi(0) &= 0^{t_1}1\,, \\ \varphi(1) &= 0^{t_2}2\,, \\ &\vdots \\ \varphi(m-2) &= 0^{t_{m-1}}(m-1)\,, \\ \varphi(m-1) &= 0^{t_m}\,, \end{aligned}$$

where $m \geq 2$, $t_1 \geq 1$, $t_m \geq 1$ and for all $i \in \{2, \ldots, m\}$,

$$t_i t_{i+1} \cdots t_m 0^\omega \prec_{\text{lex}} t_1 t_2 \cdots t_m 0^\omega\,.$$

We denote $u_n := \varphi^n(0)$ and $U_n := |u_n|$ for $n \in \mathbb{N}$.

Attractors and powers

A word u is a *power* of a word v if $u = v^k v'$, where v' is a prefix of v

Example

u=00100100100 $= (001)^3 00 = v^3 v'$ is a power of $v = 001$
kabelka $= ($*kabel*$)$*ka* is a power of *kabel*
couscous $= ($*cous*$)^2$ is a power of *cous*

Lemma

Let u be power of a word v. If Γ is an attractor of v, then $\Gamma \cup \{|v| - 1\}$ is an attractor of u.

Example

For $v = 0\mathbf{10}$ and $u = 0\mathbf{10}01001001 = (010)^3 01$, u is a power of v,
$\Gamma = \{1, 2\}$ is an attractor of v and $\Gamma \cup \{2\} = \Gamma$ is an attractor of u.

Simple Parry sequence – powers

Lemma

The word u_{n+1} is a power of u_n for all $n \in \mathbb{N}$, $n \geq m-1$.

Example

For $m = 3$ and $t_1 t_2 t_3 = 211$, $\varphi(0) = 001$, $\varphi(1) = 02$, $\varphi(2) = 0$ and the prefixes $u_n = \varphi^n(0)$ of $\mathbf{u}$ satisfy

$$u_3 = \underbrace{00100102}_{u_2}\underbrace{00100102}_{u_2}\underbrace{0010}_{u_2'} = u_2^2 u_2'$$

$$u_4 = \underbrace{00100102001001020010}_{u_3}\underbrace{00100102001001020010}_{u_3}\underbrace{00100102001}_{u_3'} = u_3^2 u_3'$$

$$u_5 = \underbrace{001001020010010200100010010200100102001000100102001}_{u_4}$$
$$\underbrace{001001020010010200100010010200100102001000100102001}_{u_4}$$
$$\underbrace{0010010200100102001000100102}_{u_4'} = u_4^2 u_4'$$

Simple Parry sequence – powers

Lemma

The word u_{n+1} is a power of u_n for all $n \geq m-1$.

Denote Q_n the length of the longest prefix of $\mathbf{u}$ that is a power of u_n

Corollary

$Q_n \geq U_{n+1}$ *for all* $n \geq m-1$.

Program

Known results for simple Parry sequence

Gheeraert, Romana, Stipulanti:

- found attractors of prefixes of size between m and $m+1$;
- searched for (minimal) attractors of prefixes being subsets of $\{U_n - 1 \ : \ n \in \mathbb{N}\}$ – *canonical attractors*

Known results for simple Parry sequence

The prefix of length ℓ of a simple Parry sequence $\mathbf{u}$ has the attractor:

length	attractor	minimal
$\ell \in [U_n, U_{n+1} - 1]$ for $0 \le n \le m - 1$	Γ_n	yes

where $\Gamma_n = \{U_0 - 1, U_1 - 1, \ldots, U_n - 1\}$ for $0 \le n \le m - 1$

Known results for simple Parry sequence

Example

For $m = 4$ and $t_1 t_2 t_3 t_4 = 2121$, $\varphi(0) = 001$, $\varphi(1) = 02$, $\varphi(2) = 003$, $\varphi(3) = 0$ and the prefixes of $\mathbf{u}$ have the following red-marked attractors

$$\begin{aligned}
u_0 &= \mathbf{0}\,, \\
u_1 &= \mathbf{001}\,, \\
v_1 &= \mathbf{0}0\mathbf{1}001\,, \\
u_2 &= \mathbf{0}0\mathbf{1}0010\mathbf{2}\,, \\
v_2 &= \mathbf{0}0\mathbf{1}0010\mathbf{2}00100102\,, \\
u_3 &= \mathbf{0}0\mathbf{1}0010\mathbf{2}00100102001 00\mathbf{3}\,, \\
v_3 &= \mathbf{0}0\mathbf{1}0010\mathbf{2}0010010200100\mathbf{3}00100102\,.
\end{aligned}$$

Known results for simple Parry sequence

The prefix of length ℓ of a simple Parry sequence $\mathbf{u}$ has the attractor:

length	attractor	minimal
$\ell \in [U_n, U_{n+1} - 1]$ for $0 \leq n \leq m-1$	Γ_n	yes
$\ell \in [U_n, U_{n+1} - 1]$ for $n \geq m$	$\Gamma_{n-1} \cup \{U_n - 1\}$	no
$\ell \in [U_n, Q_{n-1}]$ for $n \geq m$	Γ_{n-1}	yes
$\ell \in [P_n, U_{n+1}]$ for $n \geq m$	Γ_n	yes

where

$$\Gamma_n = \begin{cases} \{U_0 - 1, U_1 - 1, \ldots, U_n - 1\} & \text{for } 0 \leq n \leq m-1\,, \\ \{U_{n-m+1} - 1, U_{n-m+2} - 1, \ldots, U_n - 1\} & \text{for } n \geq m\,, \end{cases}$$

and

$$p_n = u_n u_{n-m}^{t_1 - 1} u_{n-m-1}^{t_2} \cdots u_{n-2m+1}^{t_m} \text{ is a prefix of } \mathbf{u} \text{ and } P_n = |p_n| - 1$$

If $\mathbf{P_n \leq Q_{n-1}}$, then the attractors are minimal!

Known results for simple Parry sequence

Theorem (Gheeraert, Romana, Stipulanti, 2024)

Let $\mathbf{u}$ be a simple Parry sequence with affine factor complexity, i.e., satisfying the following conditions:

1. $t_m = 1$;
2. *if there exists a word $v \neq \varepsilon$ such that v is a proper prefix and a proper suffix of $t_1 \cdots t_{m-1}$, then $t_1 \cdots t_{m-1} = w^k$ for some word w and $k \in \mathbb{N}$, $k \geq 2$.*

Then $Q_{n-1} \geq P_n$ for each $n \in \mathbb{N}$, $n \geq m$, hence

- *the prefix of $\mathbf{u}$ of length $\ell \in [U_n, P_n]$ has the attractor Γ_{n-1};*
- *the prefix of $\mathbf{u}$ of length $\ell \in [P_n, U_{n+1}]$ has the attractor Γ_n.*

Program

New results for simple Parry sequence

- Denote z_n the following prefix of a simple Parry sequence $\mathbf{u}$

$$z_n = u_n u_{n-m}^{t_1 - t_m} u_{n-m-1}^{t_2} \cdots u_{n-2m+1}^{t_m} \quad \text{and} \quad Z_n = |z_n|$$

- Compare to $p_n = u_n u_{n-m}^{t_1 - 1} u_{n-m-1}^{t_2} \cdots u_{n-2m+1}^{t_m}$

Theorem (Dvořáková, Moravcová, 2025)

Let $\mathbf{u}$ be a simple Parry sequence. Assume $t_1 > \{t_{m-1}, t_m\}$ and

$$t_i \cdots t_{m-2}(t_{m-1} + 1)0^\omega \prec_{lex} t_1 t_2 \cdots t_m 0^\omega \quad \textit{for all } i \in \{2, \ldots, m-2\}\,.$$

Then $Q_{n-1} \geq Z_n$ for each $n \in \mathbb{N}$, $n \geq m$,

- *the prefix of $\mathbf{u}$ of length $\ell \in [U_n, Z_n]$ has the attractor Γ_{n-1};*
- *the prefix of $\mathbf{u}$ of length $\ell \in [Z_n, U_{n+1}]$ has the attractor Γ_n.*

New results for simple Parry sequence

Example

For $m = 3$ and $t_1 = 3$, $t_2 = 0$, $t_3 = 2$, the morphism is defined as $\varphi(0) = 0001$, $\varphi(1) = 2$, $\varphi(2) = 00$ and the prefixes of $\mathbf{u}$ have the red-marked attractors

$$
\begin{aligned}
u_0 &= 0\,, \\
u_1 &= 0001\,, \\
u_2 &= 0001000100012\,, \\
u_3 &= \mathbf{0}00\mathbf{1}00010001\mathbf{2}000100010001200010001000 1200 \\
z_3 = u_3u_0 &= \mathbf{0}00\mathbf{1}00010001\mathbf{2}0001000100012000100010001 2000 \\
z_3 = u_3u_0 &= 000\mathbf{1}00010001\mathbf{2}00010001000120001000100012 0\mathbf{0}0
\end{aligned}
$$

New results for simple Parry sequence

Consider $m = 4$ and $t_1 t_2 t_3 t_4 = 2121$, i.e., the simplest case where neither assumptions of Theorem by Gheeraert et al. nor assumptions of Theorem by Dvořáková and Moravcová are met.
Then we find a prefix whose minimal attractor is never canonical, i.e., it is not a subset of

$$\{U_n - 1 \ : \ n \in \mathbb{N}\}.$$

New results for simple Parry sequence

Theorem 2 (Dvořáková, Moravcová, 2025)

Let **u** be a simple Parry sequence, where $t_1 > t_m$. Denote

$$k = \min\{j \in \{1, \ldots, m-1\} \ : \ t_{m-j} \neq 0\}\,.$$

For each $n \in \mathbb{N}$, $n \geq m$,

- the prefix of **u** of length $\ell \in [Z_n, U_{n+1}]$ has the attractor Γ_n;
- the prefix of **u** of length $\ell \in [U_n, Z_n]$ has the attractor

$$\Gamma = \begin{cases} \Gamma_{n-1} \cup \{U_n - U_{n-m+k} - (t_m - 1)U_{n-m} - 1\} \setminus \{U_{n-m} - 1\}\,, \\ \Gamma_{n-1} \cup \{U_n - t_m U_{n-m} - 1\} \setminus \{U_{n-m+k} - 1\}\,. \end{cases}$$

Remark

We have a similar, slightly more technical, result for $t_1 = t_m$. The attractors are again of alphabet size, thus minimal.

Binary non-simple Parry sequence – definition

Definition (Binary non-simple Parry sequence)

A *binary non-simple Parry sequence* $\mathbf{u} = \mathbf{u}(p, q)$ is the fixed point of the morphism

$$\begin{aligned} \varphi(0) &= 0^p 1\,, \\ \varphi(1) &= 0^q 1\,, \end{aligned}$$

where $p, q \geq 1$, $p > q$.

Program

New results for binary non-simple Parry sequence

Lemma

For every $n \in \mathbb{N}$,

$$\varphi^n(0) = \varphi^{n-1}(0)\varphi^{n-2}(0)\cdots\varphi(0)\mathbf{0}\cdots\mathbf{1}\varphi(1)\cdots\varphi^{n-2}(1)\varphi^{n-1}(1)\,.$$

Theorem (Dvořáková, Moravcová, 2025)

Let $\mathbf{u} = \mathbf{u}(p, q)$ be a binary non-simple Parry sequence. Then for each $n \geq 1$, the prefix $\varphi^n(0)$ has the attractor

$$\Gamma_n = \left\{ \sum_{j=0}^{n-1} |\varphi^j(0)| - 1,\ |\varphi^n(0)| - \sum_{j=1}^{n-1} |\varphi^j(1)| - 1 \right\}.$$

Techniques of the proof

- Morphic structure
- Powers
- Lexicographic conditions
- Generalized mathematical induction
- Programs in C++ (for binary, ternary and quaternary simple Parry sequences) – finding candidates for attractors, checking correctness of conjectured attractors

Program

Morphic sequences and numeration systems

Question by Gheeraert, Romana, Stipulanti:
Given a morphic sequence $\mathbf{u}$ does there exist a numeration system $\mathcal{S}$ such that $\mathbf{u}$ is $\mathcal{S}$-automatic and (minimal) attractors of the prefixes of $\mathbf{u}$ are easily described by $\mathcal{S}$?

Answer for Parry sequences:
Every Parry sequence $\mathbf{u}_\beta$ is $\mathcal{S}$-automatic for the Dumont-Thomas numeration system $\mathcal{S}$ corresponding to the morphism φ_β. Moreover, $\mathcal{S}$ coincides with the positional U-system: $U = (U_n)_{n=0}^{\infty}$ with $U_n = |\varphi^n(0)|$.

- simple Parry sequences: for each prefix, there is an attractor $\subset \{U_n - 1 \ : \ n \in \mathbb{N}\}$, not always minimal;
- binary non-simple Parry sequences: since $\varphi^n(0)$ ends in 1 for $n \geq 1$, the attractor of size two is never a subset of $\{U_n - 1 \ : \ n \in \mathbb{N}\}$

Question by Gheeraert, Romana, Stipulanti:
Given a morphic sequence $\mathbf{u}$ does there exist a numeration system $\mathcal{S}$ such that $\mathbf{u}$ is $\mathcal{S}$-automatic and (minimal) attractors of the prefixes of $\mathbf{u}$ are easily described by $\mathcal{S}$?
Answer for Parry sequences:
Every Parry sequence $\mathbf{u}_\beta$ is $\mathcal{S}$-automatic for the Dumont-Thomas numeration system $\mathcal{S}$ corresponding to the morphism φ_β. Moreover, $\mathcal{S}$ coincides with the positional U-system: $U = (U_n)_{n=0}^{\infty}$ with $U_n = |\varphi^n(0)|$.

- simple Parry sequences: for each prefix, there is an attractor $\subset \{U_n - 1 \ : \ n \in \mathbb{N}\}$, not always minimal;
- binary non-simple Parry sequences: since $\varphi^n(0)$ ends in 1 for $n \geq 1$, the attractor of size two is never a subset of $\{U_n - 1 \ : \ n \in \mathbb{N}\}$

Program

Open problems

- Gheeraert, Romana and Stipulanti studied slightly more general morphisms. It remains an open problem to find the minimal attractors of prefixes in this more general case.
- We introduced only sufficient conditions for the attractors of prefixes of simple Parry sequences to be a subset of $\{U_n - 1 \ : \ n \in \mathbb{N}\}$. It remains to find necessary conditions.

Open problems

- Gheeraert, Romana and Stipulanti studied slightly more general morphisms. It remains an open problem to find the minimal attractors of prefixes in this more general case.
- We introduced only sufficient conditions for the attractors of prefixes of simple Parry sequences to be a subset of $\{U_n - 1 \ : \ n \in \mathbb{N}\}$. It remains to find necessary conditions.
- Finding (minimal) attractors of prefixes of non-simple Parry sequences over larger alphabets

Open problems

- Gheeraert, Romana and Stipulanti studied slightly more general morphisms. It remains an open problem to find the minimal attractors of prefixes in this more general case.
- We introduced only sufficient conditions for the attractors of prefixes of simple Parry sequences to be a subset of $\{U_n - 1 \ : \ n \in \mathbb{N}\}$. It remains to find necessary conditions.
- Finding (minimal) attractors of prefixes of non-simple Parry sequences over larger alphabets
- Studying (minimal) attractors of prefixes of fixed points of morphisms in general – in which cases they are of alphabet size
  - Yes – Sturmian, episturmian, simple Parry sequences
  - No – Thue-Morse sequence

Open problems

- Gheeraert, Romana and Stipulanti studied slightly more general morphisms. It remains an open problem to find the minimal attractors of prefixes in this more general case.
- We introduced only sufficient conditions for the attractors of prefixes of simple Parry sequences to be a subset of $\{U_n - 1 \; : \; n \in \mathbb{N}\}$. It remains to find necessary conditions.
- Finding (minimal) attractors of prefixes of non-simple Parry sequences over larger alphabets
- Studying (minimal) attractors of prefixes of fixed points of morphisms in general – in which cases they are of alphabet size
  - Yes – Sturmian, episturmian, simple Parry sequences
  - No – Thue-Morse sequence

References

- F. Gheeraert, G. Romana, M. Stipulanti, *String attractors of some simple-Parry automatic sequences*, TOCS **68** (2024), 1601–1621
- L. Dvořáková, M. Moravcová, *Attractors of sequences coding β-integers*, 2025, `https://arxiv.org/abs/2511.00650`

Thank you for your attention