How to prove that a language is regular or star-free?

Jean-Éric Pin\textsuperscript{1}

\textsuperscript{1}IRIF, CNRS and Université de Paris Cité

CIRM, Mai 2022, Luminy
Table of Contents

1 Definitions and early results
2 Iteration properties
3 Rewriting systems
4 Well quasi-orders
5 Logic
6 Transductions and profinite topology
7 Functions from $\mathbb{N}$ to $\mathbb{N}$
Part V

Logic

\[ \varphi = \exists x \ a x \]

\[ L = A^* a A^* \]

\[ \varphi = \exists x \ \exists y \ \exists z \ x < y \land y < z \land a x \land b y \land c z \]

\[ L = A^* a A^* b A^* c A^* \]

\[ \varphi = \exists x \ (\forall y \ \neg y < x \land a x) \]

\[ L = a A^* \]
For each letter $a \in A$, let $a$ be a unary predicate symbol, where $ax$ is interpreted as “the letter in position $x$ is an $a$”.

To each word $u \in A^+$ is associated a structure

$$\mathcal{M}_u = (\{1, 2, \ldots, |u|\}, S, <, (a)_{a \in A})$$

where $<$ is interpreted as the usual order on $\{1, 2, \ldots, |u|\}$ and $S$ is the successor relation.

The language defined by a sentence $\varphi$ is

$$L(\varphi) = \{ u \in A^* \mid \mathcal{M}_u \text{ satisfies } \varphi \}$$
Logical fragments

Let $\text{FO}[<]$, $\text{MSO}[<]$ denote the set of first-order and monadic second-order formulas of signature $\{<, (a)_{a \in A}\}$, respectively.

Let $\text{MSO}[S]$ and $\text{SO}[S]$ denote the set of monadic second-order and second-order formulas of signature $\{S, (a)_{a \in A}\}$, respectively.
Logic on words

**Theorem** (Buchi 1960, Elgot 1961, Trakhtenbrot 1973)

\[
\text{MSO}[S] \text{ captures the class of regular languages.}
\]

**Theorem** (McNaughton 1971)

\[
\text{FO[<] captures the class of star-free languages.}
\]

In general, *second order logic* captures some non-regular languages, but two successive results led to a complete characterisation of the syntactic fragments that only capture *regular languages*. 
Existential second order (ESO or $\Sigma^1_1$)

For example, $\text{ESO}(\exists^* \forall \exists^*)$ is the class of all formulas

$$\exists R_1 \cdots \exists R_n \, \exists x_1 \cdots \exists x_p \, \forall y \, \exists z_1 \cdots \exists z_q \varphi(R_1, \ldots, R_n, x_1, \ldots, x_p, y, z_1, \ldots, z_q)$$

where $\varphi$ is quantifier-free and $R_1, \ldots, R_n$ are relation symbols.
ESO-prefix classes (pictures from EGS 2010)

Eiter, Gottlob, Gurevich 00 and Eiter, Gottlob, Schwentick 02

\[ \text{NP-tailored} \]
\[ \text{ESO}(\forall^*) \]

\[ \text{regular-tailored} \]
\[ \text{ESO}(\forall\forall) \quad \text{ESO}(\exists^*\forall\exists^*) \quad \text{ESO}(\forall\exists) \quad \text{ESO}(\exists^*\forall) \quad \text{ESO}(\forall\exists) \]

\[ \text{FO expressible (FO}(\exists^*\forall)) \]

\[ \text{regular} \quad \text{NP-hard} \]
Beyond existential second-order

For instance $\Sigma^1_2(\forall \exists)$ corresponds to formulas

$$\exists R_1 \cdots \exists R_n \ \forall S_1 \cdots \forall S_m \ \forall x \ \exists y \ \varphi(R_1, \ldots, R_n, S_1, \ldots, S_m, x, y)$$

where $R_1, \ldots, R_n, S_1, \ldots, S_m$ are relations and $\varphi$ is quantifier-free.
Linear temporal logic (LTL)

The vocabulary consists of an atomic proposition $p_a$ for each letter $a$, the usual connectives $\lor$, $\land$ and $\neg$ and the temporal operators $X$ (next), $F$ (eventually) and $U$ (until).

**Theorem (Kamp 1968)**

A language is *star-free* iff it is *LTL*-definable.
Rabin’s tree theorem

For each letter $a$, let $S_a$ be a binary relation symbol, interpreted on $A^*$ as follows: $S_a(u, v)$ iff $v = ua$.

**Theorem (Rabin 1969)**

A language is regular iff it is definable in $\text{MSO}[(S_a)_{a \in A}]$. 
Let $M$ and $N$ be monoids. A transduction $\tau: M \to N$ is a relation on $M$ and $N$, viewed as a function from $M$ to $\mathcal{P}(N)$.

The inverse transduction $\tau^{-1}: N \to M$ is defined by $\tau^{-1}(Q) = \{m \in M \mid \tau(m) \cap Q \neq \emptyset\}$.
Recognizable subsets of a monoid

Let $L$ be a subset of a monoid $M$. A monoid $N$ recognizes $L$ if there exists a surjective morphism $h : M \to N$ such that $L = h^{-1}(h(L))$.

Let $\text{Rec}(M)$ denote the set of recognizable subsets of $M$ (= recognized by some finite monoid.).

For $A^*$, recognizable $=$ regular $=$ rational.
A function \( f : M \rightarrow N \) preserves recognizability if, for each recognizable subset \( R \) of \( N \), \( f^{-1}(R) \) is recognizable.

**Proposition (Pin-Silva 2005)**

The function \( g : A^* \times \mathbb{N} \rightarrow A^* \) defined by \( g(x, n) = x^n \) preserves recognizability.

Let \( \tau_n : A^* \rightarrow (A^*)^n \) be defined by

\[
\tau_n(u) = \{(u_1, \ldots, u_n) \mid u_1 \cdots u_n = u\}
\]

Then both \( \tau_n \) and \( \tau_n^{-1} \) preserve recognizability.
Residually finite monoids

A monoid $N$ separates two elements $x, y$ of a monoid $M$ if there exists a monoid morphism $h : M \to N$ such that $h(x) \neq h(y)$.

A monoid $M$ is residually finite if any pair of distinct elements of $M$ can be separated by a finite monoid.

Let $M$ be the class of monoids that are finitely generated and residually finite. This class includes finite monoids, free monoids, free groups, trace monoids and their products.
Profinite metric on a monoid $M$ of $\mathcal{M}$

Let, for each $(u, v) \in M^2$,

$$r(u, v) = \min \{ \text{Card}(N) \mid N \text{ separates } u \text{ and } v \}$$

$$d(u, v) = 2^{-r(u,v)}$$

with the usual conventions $\min \emptyset = +\infty$ and $2^{-\infty} = 0$.

Then

$$d(u, w) \leq \max(d(u, v), d(v, w)) \quad \text{(ultrametric)}$$

$$d(uw, vw) \leq d(u, v)$$

$$d(wu, wv) \leq d(u, v)$$

Then $(M, d)$ is a metric monoid.
Uniform continuity

Proposition

Let $M, N \in \mathcal{M}$. A function $f : M \to N$ is preserves recognizability if and only if it is uniformly continuous for the profinite metrics.
Uniform continuity

Proposition

Let $M, N \in \mathcal{M}$. A function $f : M \to N$ is preserves recognizability if and only if it is uniformly continuous for the profinite metrics.

What about transductions $\tau : M \to N$?
The completion of the metric monoid \((M, d)\) is a compact metric monoid \(\widehat{M}\). The set \(\mathcal{K}(\widehat{M})\) of compact subsets of \(\widehat{M}\) is also a compact monoid for the Hausdorff metric.

The Hausdorff metric on \(\mathcal{K}(\widehat{M})\) is defined as follows. For \(K, K' \in \mathcal{K}(\widehat{M})\), let

\[
\delta(K, K') = \sup_{x \in K} d(x, K')
\]

\[
h(K, K') = \max(\delta(K, K'), \delta(K', K))
\]

+ special definition if \(K\) or \(K'\) is empty
The case of transductions

Let $M$ and $N$ be monoids of $\mathcal{M}$ and let $\tau : M \to N$ be a transduction.

Define a map $\widehat{\tau} : M \to \mathcal{K}(\widehat{N})$ by setting, for each $x \in M$, $\widehat{\tau}(x) = \tau(x)$.

Theorem (Pin-Silva 2005)

The transduction $\tau$ is preserves recognizability iff $\widehat{\tau}$ is uniformly continuous for the Hausdorff metric.
An exercise

If $L$ is regular, then

$$K = \{ u \in A^* \mid u|u| \in L \}$$

is also regular.

**Proof.** Indeed, $K = h^{-1}(L)$, where $h(u) = u|u|$. Now $h = g \circ f$

$$u \xrightarrow{f} (u, |u|) \xrightarrow{g} u|u|$$

and $f$ and $g$ are both uniformly continuous. Thus $K$ is regular.
Matrix representations

\[ \mu(a) = a \quad \mu(b) = b \quad \mu(u) = u \]

\[ f_1(u) = uu \quad f_1(u) = (\mu(u))^2 \]

\[ f_2(u) = uauu^2 \quad f_2(u) = \mu(u)\mu(u)\mu(u)^2 \]

\[ \tau_1(u) = u^* \quad \tau_1(u) = \sum_{n \geq 0} \mu(u)^n \]

\[ \tau_2(u) = \bigcup_{p \text{ prime}} u^p \quad \tau_2(u) = \sum_{p \text{ prime}} \mu(u)^p \]
\[ f(u) = a^{u|a}b^{u|b} \]

\[ \mu(a) = \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} \quad \mu(b) = \begin{pmatrix} 1 & 0 \\ 0 & b \end{pmatrix} \quad \mu(u) = \begin{pmatrix} a^{u|a} & 0 \\ 0 & b^{u|b} \end{pmatrix} \]

\[ f(u) = \mu_{1,1}(u)\mu_{2,2}(u) \]
\[ f(u) = \text{Last}(u)u \]

\[
\begin{align*}
\mu(a) &= \begin{pmatrix} a & 0 & 1 \\ 0 & a & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \quad \mu(b) &= \begin{pmatrix} b & 0 & 0 \\ 0 & b & 1 \\ 0 & 0 & 0 \end{pmatrix} \\
\mu(ua) &= \begin{pmatrix} ua & 0 & u \\ 0 & ua & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \quad \mu(ub) &= \begin{pmatrix} ub & 0 & 0 \\ 0 & ub & u \\ 0 & 0 & 0 \end{pmatrix}
\end{align*}
\]

\[ f(u) = a\mu_{1,3}(u) + b\mu_{2,3}(u) \]
Matrix representations

A transduction \( \tau : A^* \rightarrow M \) admits a matrix representation \((S, \mu)\) of degree \(n\) if there exist a monoid morphism \(\mu : A^* \rightarrow \mathcal{P}(M)^{n \times n}\) and a possibly infinite union of products \(S\) involving arbitrary subsets of \(M\) and \(n^2\) variables \(X_{1,1}, \ldots, X_{n,n}\), such that, for all \(u \in A^*\),

\[
\tau(u) = S[\mu_{1,1}(u), \ldots, \mu_{n,n}(u)].
\]

Example for \(n = 2\): Let \((P_k)_{k \geq 0}\) be subsets of \(M\).

\[
S = \bigcup_{k \in \mathbb{N}} P_0X_{1,1}^k P_kX_{2,1}X_{1,1}^k X_{2,2} P_k! X_{1,1} P_{2k}
\]
Matrix representation of transducers

Theorem (Pin-Sakarovitch 1983)

Let \((S, \mu)\) be a matrix representation of degree \(n\) of a transduction \(\tau : A^* \rightarrow M\). Let \(P\) be a subset of \(M\) recognised by a morphism \(\eta : M \rightarrow N\). Then the language \(\tau^{-1}(P)\) is recognised by the submonoid \(\eta \mu(A^*)\) of the monoid of matrices \(\mathcal{P}(N)^{n \times n}\).

Corollary

Every transduction having a matrix representation preserves recognizability.
An exercise

Let $L$ be a regular language and $T \subseteq \mathbb{N}^2$. Then

$$L_T = \{u \in A^* \mid \text{there exist } x, y \text{ and } (p, q) \in T \text{ such that } |x| = p|u|, |y| = q|u| \text{ and } xuyy \in L\}$$

is regular.

Observe that $L_T = \tau^{-1}(L)$ where the transduction

$$\tau(u) = \bigcup_{(p, q) \in T} A^p|u|uA^q|u|$$

admits the matrix representation $(S, \mu)$, with

$$\mu(u) = \begin{pmatrix}
A^{|u|} & \emptyset & \emptyset \\
\emptyset & u & \emptyset \\
\emptyset & \emptyset & A^{|u|}
\end{pmatrix}$$

and $S = \bigcup_{(p, q) \in T} X_{1,1}^pX_{2,2}X_{3,3}^q$. 
Some other examples

- **Intersection**: \( L_1 \cap L_2 = \tau^{-1}(L_1 \times L_2) \) where \( \tau : A^* \rightarrow A^* \times A^* \) is given by \( \tau(u) = u \times u \).

- **Concatenation product**: \( L_1 L_2 = \tau^{-1}(L_1 \times L_2) \) where \( \tau(u) = \{(u_1, u_2) \mid u_1 u_2 = u \} \).

  \[
  \mu(u) = \begin{pmatrix}
  (u, 1) & \{(u_1, u_2) \mid u = u_1 u_2\} \\
  \emptyset & (1, u)
  \end{pmatrix}
  \]

  (Schützenberger product)

- **Shuffle**: \( L_1 \bowtie L_2 = \tau^{-1}(L_1 \times L_2) \) where \( \tau(u) = \{(u_1, u_2) \mid u \in u_1 \bowtie u_2\} \). Here \( \mu = \tau \).

- **Union, quotients, morphisms, inverses of morphisms, and many others**
If $L$ is regular, then so are $\sqrt{L} = \{u \mid uu \in L\}$ and $\frac{1}{2}(L) = \{\text{first halves of words in } L\}$. If $L$ is star-free, then so is $\sqrt{L}$.

**Proof.** $\sqrt{L} = \tau^{-1}(L)$ where $\tau(u) = u^2$. Taking $\mu(u) = u$ shows that if a monoid recognizes $L$, then it also recognizes $\sqrt{L}$.

$\frac{1}{2}(L) = \tau^{-1}(L)$ where $\tau(u) = uA^{|u|}$. Take $\mu(u) = \begin{pmatrix} u & \emptyset \\ \emptyset & A^{|u|} \end{pmatrix}$.
Streaming string transducers

A substitution $\sigma : A^* \to B^*$ is a monoid morphism from $A^*$ to $\mathcal{P}(B^*)$.

A streaming string transducer is a sequential transducer whose outputs are substitutions.

**Theorem**

*Streaming string transducers preserve recognizability.*
An example of streaming string transducer

The function $f(a^n c b^p) = a^p b^{pn}$ can be realized by the following streaming string transducer:

$$
\sigma_1: a \mapsto a, c \mapsto c, b \mapsto b, Y \mapsto Y, X \mapsto X
$$

$$
\sigma_2: a \mapsto a, c \mapsto c, b \mapsto b, Y \mapsto Y, X \mapsto X
$$

where $A = \{a, b, c\}$, $B = A \cup \{X, Y\}$ and $\sigma, \sigma_1, \sigma_2 : B^* \rightarrow B^*$ are substitutions defined by

$X\sigma_1 = X, Y\sigma_1 = YX, d\sigma_1 = d$ for $d \in A$

$X\sigma_2 = Xb, Y\sigma_2 = Ya, d\sigma_2 = d$ for $d \in A$

$X\sigma = 1, Y\sigma = 1, d\sigma = d$ for $d \in A$
Streaming string transducers at work

The function \( f(a^n cb^p) = a^p b^{pn} \) can be realized by the following streaming string transducer:

\[
\begin{align*}
\tau(a^n cb^p) &= Y \sigma_1^n \sigma_2^p \sigma = (Y X^n) \sigma_2^p \sigma = ((Y \sigma_2^p)(X \sigma_2^p)^n) \sigma \\
&= (((Y a^p)(X b^p)^n) \sigma = a^p b^{pn}
\end{align*}
\]

\[
\begin{align*}
X \sigma_1 &= X & Y \sigma_1 &= Y X & d \sigma_1 &= d \text{ for } d \in A \\
X \sigma_2 &= X b & Y \sigma_2 &= Y a & d \sigma_2 &= d \text{ for } d \in A \\
X \sigma &= 1 & Y \sigma &= 1 & d \sigma &= d \text{ for } d \in A
\end{align*}
\]
Matrix representation of a sst

The function \( f(a^n cb^p) = a^p b^{pn} \) can be realized by the following streaming string transducer:

Let \( M \) be the monoid of all substitutions from \( B^* \) into itself under composition. Then \( \mu : A^* \to (M \cup \{0\})^{2 \times 2} \) is the morphism defined by

\[
\mu(a) = \begin{pmatrix} \sigma_1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \mu(b) = \begin{pmatrix} 0 & 0 \\ 0 & \sigma_2 \end{pmatrix}, \quad \mu(c) = \begin{pmatrix} 0 & Id \\ 0 & 0 \end{pmatrix}
\]
Effective computation

Let \( R \in \text{Rec}(B^*) \) and let \( \eta : B^* \to N \) be a monoid morphism recognizing \( R \) (\( N \) finite).

One can define a right action of \( M \) on the monoid \( \mathcal{P}(N)^B \), which induces a monoid morphism \( \pi \) from \( M \) to the monoid \( T \) of all transformations on \( \mathcal{P}(N)^B \).

**Proposition (Pin, Reynier, Villevallois, 2018)**

The language \( \tau^{-1}(R) \) is recognized by the monoid morphism \( \pi \circ \mu : A^* \to (T \cup \{0\})^{Q \times Q} \).
An other example of streaming string transducer

\[ f(u_0 \# u_1 \# u_2 \# u_3 \# u_4 \# u_5 \# \cdots \# u_n) = u_1 \# u_0 \# u_3 \# u_2 \# u_5 \# u_4 \# \cdots \# u_n \]

is realized by the sst:

\[ a \mid \sigma_1 \quad a \mid \sigma_2 \quad \# \mid \text{Id} \quad \# \mid \sigma_3 \]

where \( A = \{a, b, \#\} \), \( B = A \cup \{X, Y, Z\} \) and \( \sigma, \sigma_1, \sigma_2 : B^* \rightarrow B^* \) are substitutions defined by

\[
\begin{align*}
X \sigma_1 &= X \\
X \sigma_2 &= X \\
X \sigma_3 &= X Z \# Y \#
\end{align*}
\]

\[
\begin{align*}
Y \sigma_1 &= Ya \\
Y \sigma_2 &= Y \\
Y \sigma_3 &= 1
\end{align*}
\]

\[
\begin{align*}
Z \sigma_1 &= Z \\
Z \sigma_2 &= Za \\
Z \sigma_3 &= 1
\end{align*}
\]
Part VII

Functions from $\mathbb{N}$ to $\mathbb{N}$

Siefkes,
Decidable extensions of monadic second order successor arithmetic (1970)
Regularity-preserving functions from $\mathbb{N}$ to $\mathbb{N}$

As we have seen, the regularity-preserving functions are exactly the uniformly continuous functions from $\mathbb{N}$ to $\mathbb{N}$ for the profinite metric.

A function $f : \mathbb{N} \rightarrow \mathbb{N}$ is residually ultimately periodic (rup) if, for each monoid morphism $h$ from $\mathbb{N}$ to a finite monoid, the sequence $h(f(n))$ is ultimately periodic.

**Proposition**

A function $f : \mathbb{N} \rightarrow \mathbb{N}$ is uniformly continuous iff it is residually ultimately periodic.
Ultimately periodic functions

A function \( f : \mathbb{N} \to \mathbb{N} \) is **ultimately periodic** if there exists \( t \geq 0 \) and \( p > 0 \) such that, for all \( n \geq t \), \( f(n + p) = f(n) \). For instance, the sequence

\[
1, 4, 0, 2, 8, 1, 2, 3, 5, 2, 3, 5, 2, 3, 5, 2, 3, 5, \ldots
\]

is ultimately periodic.

A function \( f : \mathbb{N} \to \mathbb{N} \) is **ultimately periodic modulo** \( n \) if the function \( f \mod n \) is ultimately periodic. It is **cyclically ultimately periodic** (cup) if it is ultimately periodic modulo \( n \) for all \( n > 0 \).
Regularity-preserving functions from $\mathbb{N}$ to $\mathbb{N}$

**Theorem (Siefkes 1970, Seiferas-McNaughton 1976)**

A function $f : \mathbb{N} \to \mathbb{N}$ is **ultimately periodic modulo** $n$ iff for $0 \leq k < n$, the set $f^{-1}(k + n\mathbb{N})$ is **regular**.

**Theorem (Siefkes 1970, Seiferas-McNaughton 1976)**

A function $f : \mathbb{N} \to \mathbb{N}$ is **regularity-preserving** iff it is **cyclically ultimately periodic** and, for every $k \in \mathbb{N}$, the set $f^{-1}(k)$ is **regular**.
Regularity-preserving functions from $\mathbb{N}$ to $\mathbb{N}$

[Siefkes 1970]
- Every polynomial function
- $n \rightarrow 2^n$
- $n \rightarrow n!$
- $n \rightarrow 2^{2^2 \ldots ^2}$ (exponential stack of 2’s of height $n$)

[Carton-Thomas 02]
- $n \rightarrow F_n$ (Fibonacci number)
- $n \rightarrow t_n$, where $t_n$ is the prefix of length $n$ of the Prouhet-Thue-Morse sequence.
Counterexamples [Siefkes 1970]

- $n \rightarrow \lfloor \sqrt{n} \rfloor$ is not cyclically ultimately periodic and hence not regularity-preserving.

- $n \rightarrow \binom{2n}{n}$ is not ultimately periodic modulo 4 and hence not regularity-preserving. Indeed

$$\binom{2n}{n} \mod 4 = \begin{cases} 
2 & \text{if } n \text{ is a power of 2,} \\
0 & \text{otherwise.}
\end{cases}$$

Open problem?

- Is the function $n \rightarrow p_n$ regularity-preserving? ($p_n$ is the $n$-th prime number).
Application to languages

**Theorem (Seiferas, McNaughton 1976)**

Let $f : \mathbb{N} \rightarrow \mathbb{N}$ be a rup function. If $L$ is regular [star-free], then the language

$$\{ x \mid \text{there exists some } y \text{ of length } f(|x|) \text{ such that } xy \in L \}$$

is also regular [star-free].
Let $f : \mathbb{N} \rightarrow \mathbb{N}$ be a strictly increasing function. Filtering a word $u = a_0a_1 \cdots a_n$ through $f$ consists in just keeping the letters $a_i$ such that $i$ is in the range of $f$.

If $L$ is regular, is the set of words of $L$ filtered by $f$ always regular?

**Theorem** (Berstel, Boasson, Carton, Petazzoni, P. (2006))

This happens iff the function $\Delta f$ defined by $\Delta f(n) = f(n + 1) - f(n)$ is regularity-preserving.
A function \( f : \mathbb{N} \rightarrow \mathbb{N} \) is **effectively regularity-preserving** if, for each given regular subset of \( \mathbb{N} \), \( f^{-1}(R) \) is regular and effectively computable.

Recall that \( \Delta f(n) = f(n + 1) - f(n) \).

**Theorem (Carton-Thomas 02)**

Let \( \chi_P \) be the characteristic function of a predicate \( P \subseteq \mathbb{N} \). If \( \Delta \chi_P \) is effectively regularity-preserving, then the monadic second order theory \( \text{MTh}(\mathbb{N}, <, P) \) is decidable.
Closure properties of cup functions

**Theorem (Siefkes 70, Zhang 98, Carton-Thomas 02)**

Let $f, g : \mathbb{N} \to \mathbb{N}$ be cyclically ultimately periodic functions. Then so are the following functions:

1. $g \circ f$, $f + g$, $fg$, $f^g$, and $f - g$ provided that $f \geq g$ and $\lim_{n \to \infty} (f - g)(n) = +\infty$,

2. (generalised sum) $n \to \sum_{0 \leq i \leq g(n)} f(i)$,

3. (generalised product) $n \to \prod_{0 \leq i \leq g(n)} f(i)$. 
Closure properties of rup functions

Theorem

Let \( f \) and \( g \) be rup functions. Then so are \( f \circ g \), \( f + g \), \( fg \), \( f^g \), \( \sum_{0 \leq i \leq g(n)} f(i) \), \( \prod_{0 \leq i \leq g(n)} f(i) \).

Let \( f : \mathbb{N} \to \{0, 1\} \) be a non-recursive function. Then the function \( n \to (\sum_{0 \leq i \leq n} f(i))! \) is regularity-preserving but non-recursive.

Open question. Is it possible to describe the primitive recursive cup \([rup]\) functions? One could try to use a recursion scheme similar to Siefkes’ primitive recursion scheme for cup functions.
Siefkes’ recursion scheme (1970)

**Theorem**

Let $g : \mathbb{N}^k \to \mathbb{N}$ and $h : \mathbb{N}^{k+2} \to \mathbb{N}$ be cyclically ultimately periodic functions satisfying three technical conditions. Then the function $f$ defined from $g$ and $h$ by primitive recursion, i.e.

\[
\begin{align*}
    f(0, x_1, \ldots, x_k) &= g(x_1, \ldots, x_k), \\
    f(n + 1, x_1, \ldots, x_k) &= h(n, x_1, \ldots, x_k, f(n, x_1, \ldots, x_k))
\end{align*}
\]

is cyclically ultimately periodic.
The three technical conditions

(1) \( h \) is cyclically ultimately periodic in \( x_{k+2} \) of decreasing period,

(2) \( g \) is essentially increasing in \( x_k \),

(3) for all \( x \in \mathbb{N}^{k+2} \), \( x_{k+2} < h(x_1, \ldots, x_{k+2}) \).

A function \( f \) is essentially increasing in \( x_j \) iff, for all \( z \in \mathbb{N} \), there exists \( y \in \mathbb{N} \) such that for all \( x \in \mathbb{N}^n \), \( y \leq x_j \) implies \( z \leq f(x_1, \ldots, x_n) \).

A function \( f \) is c.u.p. of decreasing period in \( x_j \) iff, for all \( p \), the period of the function \( f \mod p \) in \( x_j \) is \( \leq p \).