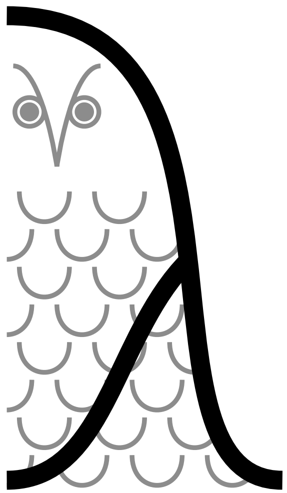
Asymptotic Approximation by Regular Languages

Ryoma Sin’ya
Akita University

YR-OWLS
30 Sep 2020



This talk is based on

[S1] Ryoma Sin'ya. Asymptotic Approximation by Regular Languages, *SOFSEM2021* (to appear), draft is available at http://www.math.akita-u.ac.jp/~ryoma

Outline

1. Motivation of this work
2. Set of natural numbers and measure density
3. Density of regular languages and REG-measurability
4. REG-(im)measurability of several languages
5. Open problems

The Primitive Words Conjecture

[Dömösi-Horvath-Ito 1991]

- A non-empty word w is said to be **primitive** if it can not be represented as a power of shorter words, i.e., $w = u^n \Rightarrow u = w$ (and $n = 1$)

  Q_A denotes the set of all primitive words over A.

- The case $\#(A) = 1$ is trivial ($\mathrm{Q}_A = A$). Here after we only consider the case $A = \{a, b\}$ for Q_A, and simply write Q.

Example： $ababa \in \mathrm{Q}$ $ababab = (ab)^3 \notin \mathrm{Q}$

Conjecture: Q is not context-free.

Why is “primitivity” important?

- Primitive words are like prime numbers.
  Fact: For every non-empty word w, there exists a unique primitive word v such that $w = v^k$ for some $k \geq 1$.

- For a word $w = uv$, we denote its *conjugate* (by u) vu by $u^{-1}wu = vu$.
  If u and v are non-empty, $u^{-1}wu$ is called a *proper* conjugate.
  Fact: w is primitive ⇔ $w \neq u^{-1}wu$ for every proper conjugate.

  Note: if we regard a conjugation as a (partial) morphism on words, “w is primitive” means “w has no non-trivial automorphism” (cf. rigid graphs, rigid models in model theory) .

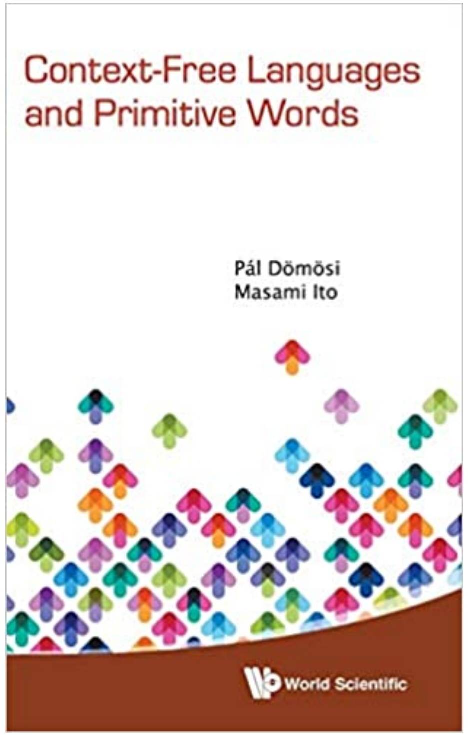
- Primitive words and its special class called *Lyndon words* play a central role in algebraic coding theory and combinatorics on words, also in text compression (cf. Lyndon factorisation, Burrows–Wheeler transformation).

The Primitive Words Conjecture

[Dömösi-Horvath-Ito 1991] **On the Connection between Formal Languages and Primitive Words**

Masami Ito

Pál Dömösi



[Dömösi-Ito 2014]

The Primitive Words Conjecture

[Dömösi-Horvath-Ito 1991] **On the Connection between Formal Languages and Primitive Words**

Masami Ito

Pál Dömösi

Szilárd Fazekas

My motivating intuition

(Intuition 1) Q is "very large" while there is no "good approximation"
by regular languages.

(Intuition 2) Every "very large" context-free language has some
"good approximation" by regular languages.

My (naive) idea: if we can formalise the above intuition and prove it, then the primitive words conjecture is true!

→ I proved that (the formal statement) of Intuition 1 is true, but Intuition 2 is false.

Approximation of languages

We adopt and extend Buck's *measure density*
to formalise "approximation by regular languages".

- Measure density [Buck 1946]

- Rough set approximation [Păun-Polkowski-Skowron 1996]
- Minimal cover-automata [Câmpeanu-Sânten-Yu 1999]
- Minimal regular cover [Domaratzki-Shallit-Yu 2001]
- Convergent-reliability / Slender-reliability [Kappes-Kintala 2004]
- Bounded-ε-approximation [Eisman-Ravikumar 2005]
- Degree of approximation [Cordy-Salomaa 2007]

Outline

1. Motivation of this work
2. Set of natural numbers and measure density
3. Density of regular languages and REG-measurability
4. REG-(im)measurability of several languages
5. Open problems

Natural density of a subset of $\mathbb{N}$ $(\ni 0)$

- For an arithmetic progression

$$S = \{cn + d \mid n \in \mathbb{N}\}$$

we define its *natural density* $\delta(S)$ as

- if $c = 0$ (i.e., $S = \{d\}$) then $\delta(S) = 0$
- if $c \neq 0$ (i.e., S is infinite) then $\delta(S) = \dfrac{1}{c}$

Intuitively, $\delta(S)$ represents the "largeness" of S. More formally, it represents the **probability** that a randomly chosen natural number n is in S.

Measure density of a subset of $\mathbb{N}$

[Buck 1946] "The measure theoretic approach to density"

- For a set of numbers $S \subseteq \mathbb{N}$, its *outer measure* $\mu^*(S)$ of S is defined as

$$\mu^*(S) = \inf\left\{ \sum_i \delta(X_i) \mid S \subseteq X, X \text{ is a disjoint union of finitely many arithmetic progressions } X_1, \ldots, X_k \right\}$$

- If a set $S \subseteq \mathbb{N}$ satisfies the condition

$$\mu^*(S) + \mu^*(\overline{S}) = 1 \qquad (☆)$$

then we call $\mu^*(S)$ *the measure density of* S, and we say that "S is *measurable*".

Theorem (Buck)：

$$\mathscr{D}_0 \subsetneq \mathscr{D}_\mu$$

- The class $\mathscr{D}_\mu$ of all subsets of $\mathbb{N}$ satisfying (☆) is the *Carathéodory extension* of

$$\mathscr{D}_0 = \left\{ X \subseteq \mathbb{N} \mid X \text{ is a disjoint union of finitely many arithmetic progresssions} \right\}$$

Observation

- $\mathscr{D}_0 = \left\{X \subseteq \mathbb{N} \mid X \text{ is a finitely many disjoint union of arithemtic progressions}\right\}$ can be seen as the class REG_A of regular languages over a unary alphabet $A = \{a\}$:

$$\mathscr{D}_0 = \{\{\,|w|\mid w \in L\} \mid L \in \mathrm{REG}_A\}$$

The set of lengths of words in a regular language L (i.e., the *Parikh image of* L) is a finite union of arithmetic progressions (i.e., *ultimately periodic set*).

If we can define a "density" notion on REG_A for an arbitrary alphabet A, we can naturally extend Buck's measure density to formal languages!

Outline

1. Motivation of this work
2. Set of natural numbers and measure density
3. <u>Density of regular languages and REG-measurability</u>
4. REG-(im)measurability of several languages
5. Open problems

Density of formal languages

- The *asymptotic density* $\delta_A(L)$ of a language L over A is defined as

$$\delta_A(L) = \lim_{n\to\infty} \frac{\#(L \cap A^n)}{\#(A^n)}$$

- The *density* $\delta_A^*(L)$ is defined as

$$\delta_A^*(L) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \frac{\#(L \cap A^i)}{\#(A^i)}$$

Fact: if $\delta_A(L)$ converges then $\delta_A^*(L)$ also converges, and moreover $\delta_A(L) = \delta_A^*(L)$.

But the converse is not true!

trivial example: $L = (AA)^*$

$\delta_A(L) = \bot$ (diverges) but

$\delta_A^*(L) = 1/2$

Density of formal languages

- The *asymptotic density* $\delta_A(L)$ of a language L over A is defined as

$$\delta_A(L) = \lim_{n\to\infty} \frac{\#(L \cap A^n)}{\#(A^n)}$$

- The *density* $\delta^*_A(L)$ is defined as

$$\delta^*_A(L) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \frac{\#(L \cap A^i)}{\#(A^i)}$$

Fact1 (cf. [Salomaa-Soittla 1978]): for any regular language L over A, $\delta^*_A(L)$ converges to a rational number.

Fact2 (cf. [S2]): A regular language L is *not null* (i.e., $\delta^*_A(L) \neq 0$) if and only if L is *dense* (i.e., $L \cap A^*wA^* \neq \emptyset$ for any $w \in A^*$).

Not null: measure theoretic "largeness"
Dense: topological "largeness"

Note: "L is not null $\Rightarrow$ L is dense" is true for any language L, but
"L is dense $\Rightarrow$ L is not null" is false for general non-regular languages.

Density of formal languages

Note: "L is not null $\Rightarrow$ L is dense" is true for any language L, but
"L is dense $\Rightarrow$ L is not null" is false for general non-regular languages.

Infinite Monkey Theorem (cf. [Borel 1913]): $\delta_A(A^*wA^*) = 1$ for any $w \in A^*$.

L is not dense means that there exists w such that $L \cap A^*wA^* = \varnothing$
(such word is called a *forbidden word* of L),
thus $\delta_A(L) \leq 1 - \delta_A(A^*wA^*) = 0$ by the infinite monkey theorem.

The *semi-Dyck* language $\mathrm{D} = \{\varepsilon, (), (()), ()(), ((())), \ldots\}$ over $A = \{(,)\}$
is dense, but actually null.

$($ $)(()($ $))$

Density of formal languages

- The *asymptotic density* $\delta_A(L)$ of a language L over A is defined as

$$\delta_A(L) = \lim_{n\to\infty} \frac{\#(L \cap A^n)}{\#(A^n)}$$

- The *density* $\delta_A^*(L)$ is defined as

$$\delta_A^*(L) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \frac{\#(L \cap A^i)}{\#(A^i)}$$

Fact1 (cf. [Salomaa-Soittla 1978]): for any regular language L over A, $\delta_A^*(L)$ converges to a rational number.

Fact2 (cf. [S2]): A regular language L is *not null* (i.e., $\delta_A^*(L) \neq 0$) if and only if L is *dense* (i.e., $\forall w \in A^*\ L \cap A^*wA^* \neq \emptyset$).

Measure density of languages

- We now consider the Carathéodory extension of the class of regular languages:

For $L \subseteq A^*$, its *outer measure* is defined as
$\overline{\mu}_{\mathrm{REG}}(L) = \inf\{\delta_A^*(R) \mid L \subseteq R \in \mathrm{REG}_A\}$.

We say that L is *REG-measurable* if $\overline{\mu}_{\mathrm{REG}}(L) + \overline{\mu}_{\mathrm{REG}}(\overline{L}) = 1$ holds.

Lemma: the followings are equivalent

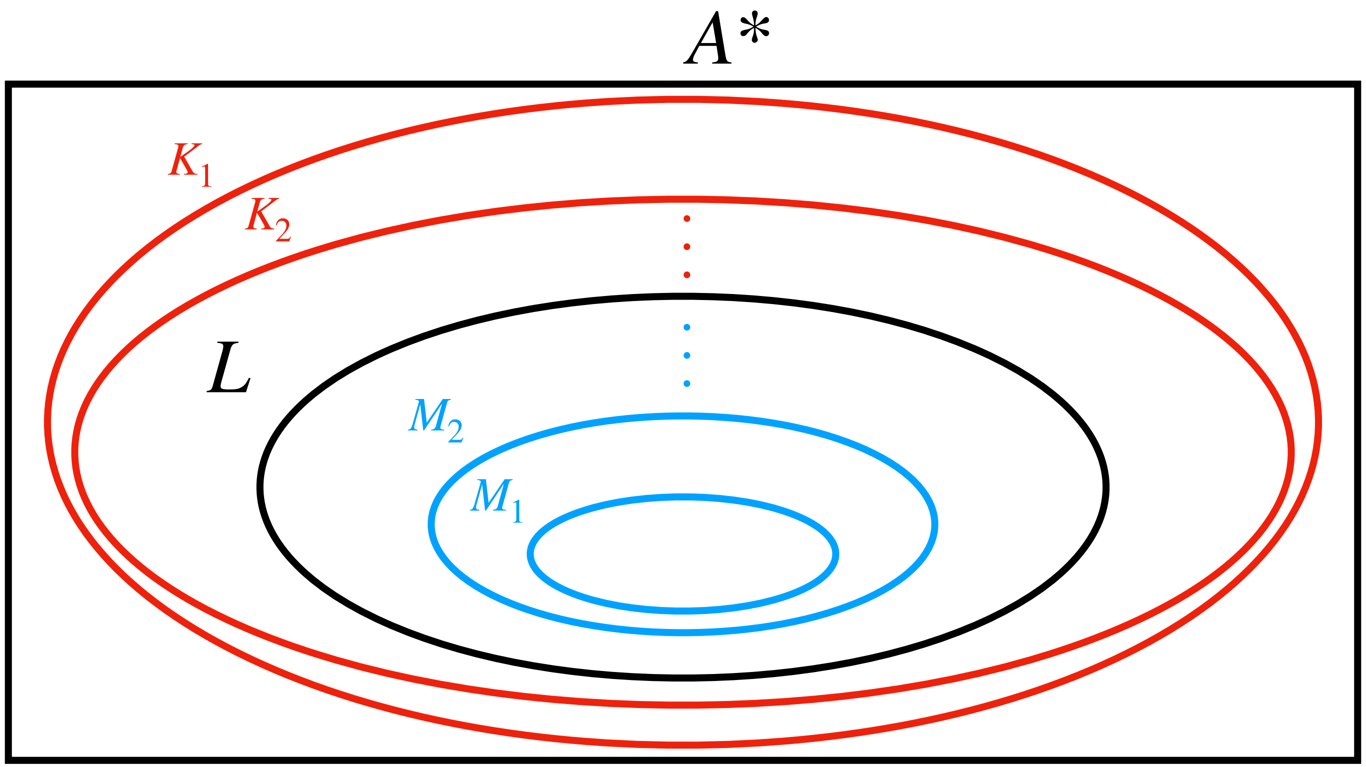
(1) L is REG-measurable

(2) $\overline{\mu}_{\mathrm{REG}}(L) = \underline{\mu}_{\mathrm{REG}}(L) = \sup\{\delta_A^*(R) \mid L \supseteq R \in \mathrm{REG}_A\}$

the *inner measure* of L

Note: $\underline{\mu}_{\mathrm{REG}}(L) \leq \delta_A^*(L) \leq \overline{\mu}_{\mathrm{REG}}(L)$ always holds (if $\delta_A^*(L)$ is defined).

Measure density of languages



L is REG-measurable if we can take an infinite sequence of pairs or regular languages $(M_n \subseteq L \subseteq K_n)_n$ such that $\lim_{n\to\infty} \delta_A^*(K_n \backslash M_n) = 0$.

Outline

1. Motivation of this work
2. Set of natural numbers and measure density
3. Density of regular languages and REG-measurability
4. <u>REG-(im)measurability of several languages</u>
5. Open problems

Example of REG-measurable CFLs

Theorem:

The semi-Dyck language $\mathsf{D} = \{\varepsilon, ab, aabb, abab, \ldots\}$ over $A = \{a, b\}$ is REG-measurable.

Note: D is null, but there does not exist a null regular superset $\mathsf{D} \subseteq L$.

(D is dense implies $\mathsf{D} \subseteq L$ is dense, and thus L is not null by Fact2)

Proof: Let $L_k = \{w \in A^* \mid |w|_a = |w|_b \mod k\}$ for each $k \geq 1$.

the # of occurrences of a in w

Then, for each $k \geq 1$, $\mathsf{D} \subseteq L_k$ and $\delta_A^*(L_k) = \dfrac{1}{k} \to 0$ (if $k \to \infty$) .

Thus the infinite sequence $(\emptyset, L_k)_{k \geq 1}$ converges to D .

Example of REG-measurable CFLs

Theorem: The following languages are all REG-measurable.

1. $\mathrm{O}_3 = \{w \in \{a, b, c\}^* \mid |w|_a = |w|_b$ or $|w|_a = |w|_c\}$
2. $\mathrm{O}_4 = \{w \in \{x, \bar{x}, y, \bar{y}\}^* \mid |w|_x = |w|_{\bar{x}}$ or $|w|_y = |w|_{\bar{y}}\}$
3. $\mathrm{P} = \{w \in \{a, b\}^* \mid w = \mathrm{reverse}(w)\}$ (the set of all *palindromes*)
4. $\mathrm{G} = \{a^{n_1}ba^{n_2}b\cdots a^{n_k}b \mid k \geq 1, n_i \neq i$ for some $i\}$ (the *Goldstine* language)

Note:

(1) and (2) are inherently ambiguous context-free languages [Flajolet 1985].

The generating function of (4) is transcendental (i.e., not algebraic) [Flajolet 1987], thus (4) is also inherently ambiguous by Chomsky-Schützenberger theorem.

Example of REG-measurable CFLs

Theorem: The following languages are all REG-measurable.

1. $\mathsf{O}_3 = \{w \in \{a,b,c\}^* \mid |w|_a = |w|_b$ or $|w|_a = |w|_c\}$
2. $\mathsf{O}_4 = \{w \in \{x,\bar{x},y,\bar{y}\}^* \mid |w|_x = |w|_{\bar{x}}$ or $|w|_y = |w|_{\bar{y}}\}$
3. $\mathsf{P} = \{w \in \{a,b\}^* \mid w = \mathrm{reverse}(w)\}$ (the set of all *palindromes*)
4. $\mathsf{G} = \{a^{n_1}ba^{n2}b\cdots a^{n_k}b \mid k \geq 1, n_i \neq i$ for some $i\}$ (the *Goldstine* language)
5. $\mathsf{K} = S_1\{c\}A^* \cup S_2\{c\}A^*$ where $A = \{a,b,c\}$,
   $S_1 = \{a\}\{b^ia^i \mid i \geq 1\}^*$ and $S_2 = \{a^ib^{2i} \mid i \geq 1\}^*\{a\}^+$.

Note: the density of (5) is transcendental [Kemp 1980], thus it is inherently ambiguous by the fact [Berstel 1972] that the density of every unambiguous context-free language is algebraic.

Example of REG-measurable CFLs

Theorem:

For every alphabet A and a language $L \subseteq A$, its *suffix extension by* $c \notin A$

$L' = L\{c\}(A \cup \{c\})^*$ is REG-measurable.

Corollary: $\mathrm{K} = (S_1 \cup S_2)\{c\}A^*$ is REG-measurable (because $S_1, S_2 \subseteq A \backslash \{c\}$).

Corollary: There exist **uncountably many** REG-measurable languages.

REG-gap: complexity of immeasurable sets

- For a language $L \subseteq A^*$ the difference $\overline{\mu}_{\mathrm{REG}}(L) - \underline{\mu}_{\mathrm{REG}}(L)$ of outer and inner measure is called the *REG-gap* of L.

REG-gap represents how a given language is “hard to approximate”.

(Intuition 1) Q is “very large” while there is no “good approximation” by regular languages.

Formal statement: Q is co-null (i.e., $\delta_A^*(\mathrm{Q}) = 1$) but $\underline{\mu}_{\mathrm{REG}}(\mathrm{Q}) = 0$.

(Intuition 2) Every “very large” context-free language has some “good approximation” by regular languages.

Formal statement: Every co-null context-free language L satisfies $\underline{\mu}_{\mathrm{REG}}(L) > 0$.

REG-immesurability of Q

(Intuition 1) Q is "very large" while there is no "good approximation"
by regular languages.

Formal statement: Q is co-null (i.e., $\delta_A^*(\mathrm{Q}) = 1$) but $\underline{\mu}_{\mathrm{REG}}(\mathrm{Q}) = 0$.

Theorem (1): Q is co-null.

Theorem (2): Every regular subset of Q is null. In particular, every non-null regular language contains infinitely many non-primitive words.

Note: The proof of Theorem (2) uses basic semigroup theory
(Green's relation and Green's theorem)

REG-immesurability of context-free langugaes

(Intuition 2) Every "very large" context-free language has some "good approximation" by regular languages.

Formal statement: Every co-null context-free language L satisfies $\underline{\mu}_{\mathrm{REG}}(L) > 0$.

Theorem: A deterministic context-free language $\mathsf{M}_2 = \{w \in \{a,b\}^* \mid |w|_a > 2|w|_b\}$ over $A = \{a,b\}$ is null but $\overline{\mu}_{\mathrm{REG}}(\mathsf{M}_2) = 1$, i.e., whose REG-gap is 1.

Corollary: $\overline{\mathsf{M}_2}$ is co-null (deterministic) context-free language with $\underline{\mu}_{\mathrm{REG}}(\overline{\mathsf{M}_2}) = 0$.

Note: This counter-example is inspired by a result of [Eisman-Ravikumar 2011]. They showed that the *majority language* $\mathsf{M} = \{w \in \{a,b\}^* \mid |w|_a > |w|_b\}$ is "hard to approximate".

REG-immesurability of context-free langugaes

Theorem: A deterministic context-free language

$\mathsf{M}_2 = \{w \in \{a,b\}^* \mid |w|_a > 2|w|_b\}$ over $A = \{a,b\}$ is null

but $\overline{\mu}_{\mathsf{REG}}(\mathsf{M}_2) = 1$, i.e., whose REG-gap is 1.

Proof: $\delta_A^*(\mathsf{M}_2) = 0$ can be shown by using the law of large numbers.

For a regular language L with $\delta_A^*(L) < 1$, we show that $\mathsf{M}_2 \subsetneq L$ (i.e., $\overline{L} \cap \mathsf{M}_2 \neq \emptyset$).

Let $\eta : A^* \to M = A^*/\simeq_{\overline{L}}$ be the syntactic morphism of $\overline{L}$.

$$c = \max_{m \in M} \min_{w \in \eta^{-1}(m)} |w| \qquad a^{4c+1}$$

$\overline{L}$ is non-null implies $\overline{L}$ is dense (infinite monkey theorem)

$\exists x, y$ such that $|x|, |y| \leq c$ and $xa^{4c+1}y \in \overline{L}$

$$|xa^{4c+1}y|_b \leq |x| + |y| \leq 2c < \frac{1}{2}|xa^{4c+1}y|_a$$

Thus $xa^{4c+1}y \in \mathsf{M}_2$ and $\mathsf{M}_2 \subsetneq L$

REG-immesurability of context-free langugaes

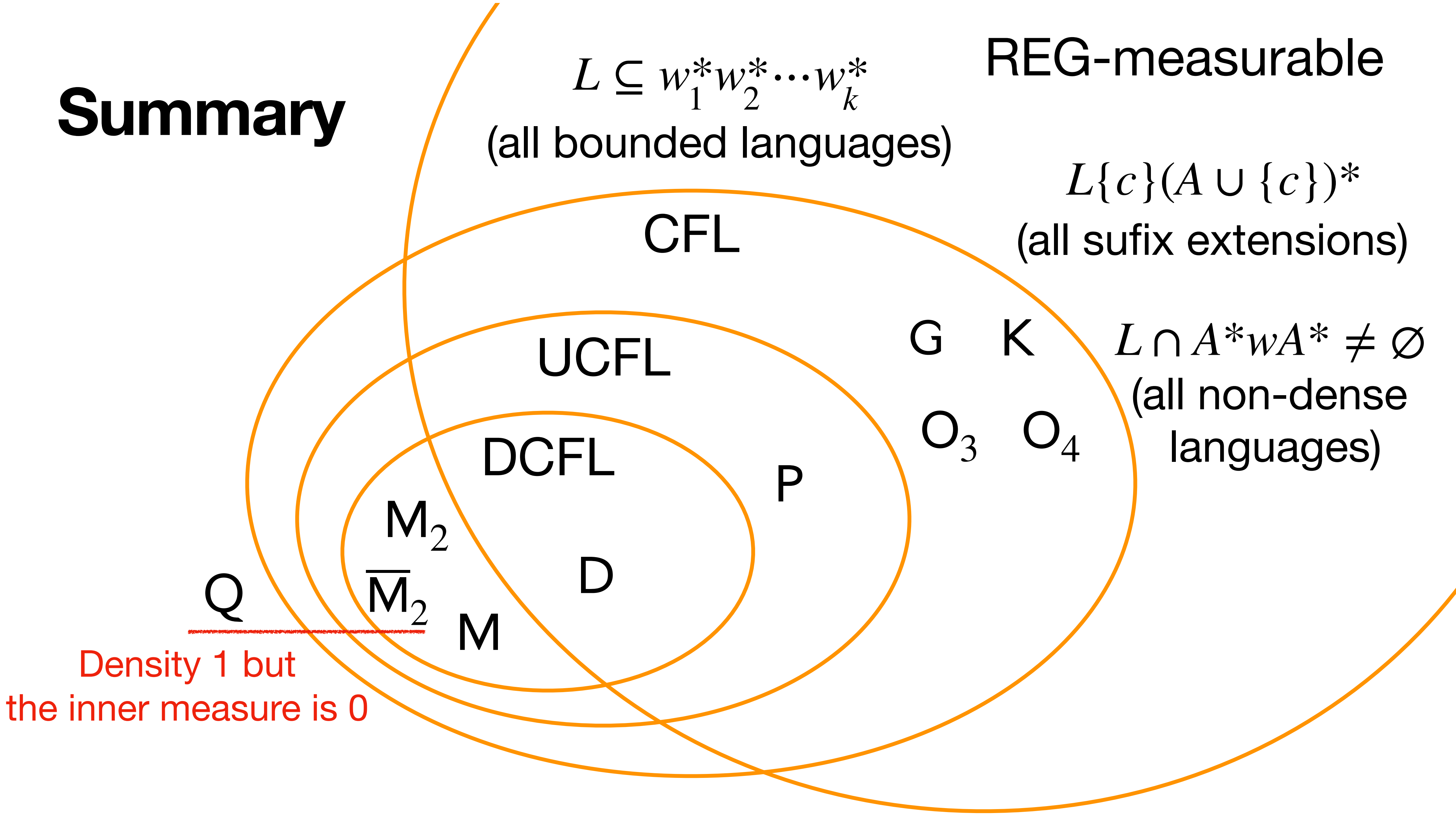
(Intuition 2) Every "very large" context-free language has some "good approximation" by regular languages.

Formal statement: Every co-null context-free language L satisfies $\underline{\mu}_{\mathrm{REG}}(L) > 0$.

Theorem: A deterministic context-free language $\mathsf{M}_2 = \{w \in \{a,b\}^* \mid |w|_a > 2|w|_b\}$ over $A = \{a,b\}$ is null but $\overline{\mu}_{\mathsf{REG}}(\mathsf{M}_2) = 1$, i.e., whose REG-gap is 1.

Corollary: $\overline{\mathsf{M}_2}$ is co-null (deterministic) context-free language with $\underline{\mu}_{\mathsf{REG}}(\overline{\mathsf{M}_2}) = 0$.

Summary



$L \subseteq w_1^* w_2^* \cdots w_k^*$
(all bounded languages)
REG-measurable
$L\{c\}(A \cup \{c\})^*$
(all sufix extensions)
CFL
UCFL
DCFL
G
K
$L \cap A^*wA^* \neq \emptyset$
(all non-dense languages)
O_3
O_4
P
M_2
D
Q
$\overline{M}_2$
M
Density 1 but
the inner measure is 0

Outline

1. Motivation of this work
2. Set of natural numbers and measure density
3. Density of regular languages and REG-measurability
4. REG-(im)measurability of several languages
5. Open problems

Open problems

1. Can we give an alternative characterisation of the class of null (resp. co-null) context-free languages?

   Note: it is **undecidable** whether a given CFG generates null (resp. co-null) CFL
   [Nakamura 2019].

2. Can we give an alternative characterisation of REG-measurable (context-free) languages?

   Note: it is **undecidable** whether a given CFG generates REG-measurable CFL,
   because REG-measurability is preserved under left/right quotients
   thus we can apply Greibach's metatheorem.

Open problems

3. Can we find a language class that “separates” Q and CFLs? i.e., is there a language class $\mathscr{C}$ such that
   - Q has full $\mathscr{C}$-gap but no co-null context-free language has full $\mathscr{C}$-gap, or
   - Q is $\mathscr{C}$-immeasurable but every co-null context-free language is $\mathscr{C}$-measurable?

Note: measurability can be parameterised by a language class $\mathscr{C}$:

Define the outer measure of L over A as

$$\overline{\mu}_{\mathscr{C}} = \{\delta_A^*(K) \mid L \subseteq K \in \mathscr{C}\}$$

and L is said to be $\mathscr{C}$-measurable if $\overline{\mu}_{\mathscr{C}}(L) + \overline{\mu}_{\mathscr{C}}(\overline{L}) = 1$.

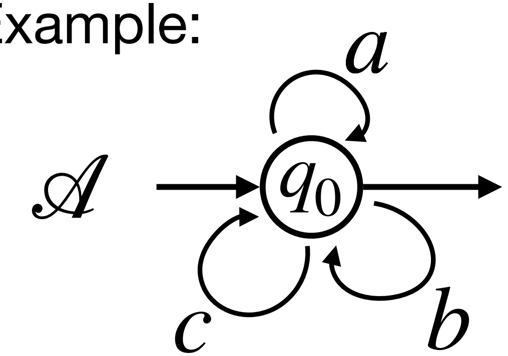
What’s happen if we consider $\mathscr{C} =$ DCFL, UCFL, CFL or UnCA?

Digression: constrained automata

- A constrained automaton is a pair $(\mathscr{A}, S)$ of a finite automaton $\mathscr{A}$ and a semi-linear set $S \subseteq \mathbb{N}^d$ whose dimension d is the # of transition rules of $\mathscr{A}$.
(i.e., Presburger definable set)

$(\mathscr{A}, S)$ accepts a word w iff there exists an accepting run ρ labeled by w and the vector $(n_1, n_2, \ldots, n_d)$ is in S where n_i is the number of occurrences the i-th transition rule in ρ.

Example:



$L((\mathscr{A}, S)) = \text{MIX} = \{w \in \{a, b, c\}^* \mid |w|_a = |w|_b = |w|_c\}$
where $S = \{(n, n, n) \mid n \in \mathbb{N}\}$.

Digression: constrained automata

- The class of **unambiguous** constrained automata is a very well-behaved class:
  - Many counting-type languages (including MIX, O_3, O_4, M and $\overline{\mathrm{M}}_2$) are in UnCA (UnCA = the class of unambiguous constrained automata recognisable languages).
  - Every UnCA language has a holonomic generating function (cf. [Bostan et al. 2020]).
  - UnCA is closed under Boolean operations and quotients [Cadilhac et al. 2012].
  - The regularity for UnCA is decidable [Cadilhac et al. 2012].
  - The context-freeness for some subclass of UnCA is decidable [S3].

Open problems

1. Can we give an alternative characterisation of the class of null (resp. co-null) context-free languages?

2. Can we give an alternative characterisation of REG-measurable (context-free) languages?

3. Can we find a language class that "separates" Q and CFLs? i.e., is there a language class $\mathscr{C}$ such that
   - Q has full $\mathscr{C}$-gap but no co-null context-free language has full $\mathscr{C}$-gap, or
   - Q is $\mathscr{C}$-immeasurable but every co-null context-free language is $\mathscr{C}$-measurable?

(Akita-Inu)

References (approximation)

- [Buck 1946] The measure theoretic approach to density, *AJM*.
- [Eisman-Ravikumar 2005] Approximate recognition of non-regular languages by finite automata, *ACSC2005*.
- [Câmpeanu-Sânten-Yu 1999] Minimal cover-automata for finite languages, *TCS*.
- [Cordy-Salomaa 2007] On the existence of regular approximations, *TCS*.
- [Domaratzki-Shallit-Yu 2001] Minimal covers of formal languages, *DLT2001*.
- [Păun-Polkowski-Skowron 1996] Rough-Set-Like Approximations of Context-Free and Regular, *IPMU1996*.
- [Kappes-Kintala 2004] Tradeoffs between reliability and conciseness 570 of deterministic finite automata, *JALC*.

References (density, ambiguity, etc.)

- [Berstel 1972] Sur la densité asymptotique de langages formels, *ICALP1972*.
- [Borel 1972] Mécanique Statistique et Irréversibilité, *J. Phys*.
- [Bostan et al. 2020] Weakly-Unambiguous Parikh Automata and Their Link to Holonomic Series, *ICALP2020*.
- [Cadilhac et al. 2012] Unambiguous Constrained Automata, *DLT2012*.
- [Dömösi-Ito 2014] Context-Free Languages And Primitive Words.
- [Dömösi-Horvath-Ito 1991] On the Connection between Formal Languages and Primitive Words.
- [Flajolet 1985] Ambiguity and transcendence, *ICALP1985*.
- [Flajolet 1987] Analytic models and ambiguity of context-free languages, *TCS*.
- [Kemp 1980] A note on the density of inherently ambiguous context-free languages, *Acta Informatica*.
- [Nakamura 2019] Computational Complexity of Several Extensions of Kleene Algebra, *Ph.D. Thesis* (Tokyo Tech).
- [Salomaa-Soittla 1978] Automata Theoretic Aspects of Formal Power Series.

References (my work)

- [S1] Asymptotic Approximation by Regular Languages, *SOFSEM2021* (to appear).
- [S2] An Automata Theoretic Approach to the Zero-One Law for Regular Languages, *GandALF2015*.
- [S3] Context-Freeness of Word-MIX Languages, *DLT2020*.

The full versions are all available at http://www.math.akita-u.ac.jp/~ryoma