Asymptotic Approximation in Formal Languages (CLA 2023 talk abstract)

Ryoma Sin'ya

Akita University (ryoma@math.akita-u.ac.jp)

$\mathcal{C}$-measurability for a class $\mathcal{C}$ of languages is introduced by [1] and it was used for classifying non-regular languages by using regular languages. A language L is said to be $\mathcal{C}$-measurable if there is an infinite sequence of languages in $\mathcal{C}$ that converges to L. Roughly speaking, L is $\mathcal{C}$-measurable means that it can be approximated by a language in $\mathcal{C}$ with *arbitrary high precision*: the notion of "precision" is formally defined by the *density* δ of formal languages:

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

Hence that a language L is not $\mathcal{C}$-measurable ($\mathcal{C}$-*im*measurable) means that L has a complex shape so that it can not be approximated by languages in $\mathcal{C}$. For example, in [1] it is shown that, while many context-free languages are REG-measurable (where REG is the class of all regular languages), some simple deterministic context-free languages and the set of all *primitive words* are REG-*im*measurable. While the membership problem for a given language L and a class $\mathcal{C}$ just asks whether $L \in \mathcal{C}$, the $\mathcal{C}$-measurability asks the existence of an infinite sequence of languages in $\mathcal{C}$ that converges to L. In this sense, measurability is much more difficult than the membership problem and its analysis is a challenging task. For example, the author [2] showed that, for the class SF of all star-free languages, the class of all SF-measurable regular languages strictly contains SF but does not contain some regular languages. However, the decidability of SF-measurability for regular languages is still unknown. For some restricted subclasses $\mathcal{C}$ of star-free languages, the decidability of $\mathcal{C}$-measurability is known [4,3].

In this talk, we give a brief overview of the known decidability results on $\mathcal{C}$-measurability, and introduce several open problems and future directions.

References

1. Sin'ya, R.: Asymptotic approximation by regular languages. In: Current Trends in Theory and Practice of Computer Science. pp. 74–88 (2021)
2. Sin'ya, R.: Carathéodory extensions of subclasses of regular languages. In: Developments in Language Theory. pp. 355–367 (2021)
3. Sin'ya, R.: Measuring power of generalised definite languages. In: Implementation and Application of Automata. pp. 278–289. Springer International Publishing (2023)
4. Sin'ya, R., Yamaguchi, Y., Nakamura, Y.: Regular languages that can be approximated by testing subword occurrences. Computer Software **40**(2), 49–60 (2023), (written in Japanese)