ON THE COMPLEXITY OF LEARNING PROGRAMS

VASCO BRATTKA

Abstract. Given a computable sequence of natural numbers, it is a natural task to find a Gödel number of a program that generates this sequence. It is easy to see that this problem is neither continuous nor computable. In algorithmic learning theory this problem is well studied from several perspectives and one question studied there is for which sequences this problem is at least learnable in the limit. Here we study the problem on all computable sequences and we classify the Weihrauch complexity of it. For this purpose we can, among other methods, utilize the amalgamation technique known from learning theory. As a benchmark for the classification we use closed and compact choice problems and their jumps on natural numbers, which correspond to induction and boundedness principles, as they are known from the Kirby-Paris hierarchy in reverse mathematics. We provide a topological as well as a computability-theoretic classification, which reveal some significant differences.

1. Summary

We study the Weihrauch complexity of the problem of finding a Gödel number of a given computable sequence. We formalize the underlying problem as follows. The *Gödel (numbering) problem* $\mathsf{G} :\subseteq \mathbb{N}^\mathbb{N} \rightrightarrows \mathbb{N}$ can be defined by

$$\mathsf{G} :\subseteq \mathbb{N}^\mathbb{N} \rightrightarrows \mathbb{N}, p \mapsto \{i \in \mathbb{N} : \varphi_i = p\},$$

where the domain $\mathrm{dom}(\mathsf{G})$ consists of all computable sequences and $\varphi : \mathbb{N} \to \mathcal{P}$ denotes some standard Gödel numbering of the set of computable sequences $\mathcal{P}$.

Another perspective one could take is to ask what additional useful information is carried by a program i that the sequence p itself does not make accessible? This question has been studied by Hoyrup and Rojas [HR17] and their answer could be summarized briefly as follows:

Slogan 1.0.1 (Hoyrup and Rojas 2017)**.** The only useful additional information carried by a program compared to the natural number sequence it represents, is an upper bound on the Kolmogorov complexity of the sequence.

One of our goals is to study in which sense this slogan can be converted into theorems on the Weihrauch complexity of the Gödel problem. For this purpose we introduce a number of further related problems. For our study the *Kolmogorv complexity* is the function $\mathsf{K} :\subseteq \mathbb{N}^\mathbb{N} \rightrightarrows \mathbb{N}$ with $\mathrm{dom}(\mathsf{K}) = \mathrm{dom}(\mathsf{G})$ defined by

$$\mathsf{K} :\subseteq \mathbb{N}^\mathbb{N} \to \mathbb{N}, p \mapsto \min \mathsf{G}(p)$$

for all $p \in \mathrm{dom}(\mathsf{K})$. The study by Hoyrup and Rojas also motivates to investigate the following variant of G

$$\mathsf{G}_\geq :\subseteq \mathbb{N}^\mathbb{N} \times \mathbb{N} \rightrightarrows \mathbb{N}, (p, m) \mapsto \mathsf{G}(p)$$

with $\mathrm{dom}(\mathsf{G}_\geq) := \{(p, m) \in \mathbb{N}^\mathbb{N} \times \mathbb{N} : m \geq \mathsf{K}(p)\}$. That is $\mathsf{G}_\geq$ is the Gödel problem that gets an upper bound on the Kolmogorov complexity as additional input information. We note that the output does not need to satisfy the input bound m according to this definition. Yet another function that one can consider is

$$\mathsf{K}_\geq : \mathbb{N}^\mathbb{N} \rightrightarrows \mathbb{N}, p \mapsto \{m \in \mathbb{N} : m \geq \mathsf{K}(p)\}$$

that just yields an upper bound on the Kolmogorov complexity of the input sequence, again with $\mathrm{dom}(\mathsf{K}_{\geq}) := \mathrm{dom}(\mathsf{G})$.

Our goal is to classify the Weihrauch complexity and also its topological counterpart of the problems mentioned above. The way we will calibrate this complexity is with the help of the problems

$$\mathsf{K}_{\mathbb{N}} <_{\mathrm{W}} \mathsf{C}_{\mathbb{N}} <_{\mathrm{W}} \mathsf{K}_{\mathbb{N}}' <_{\mathrm{W}} \mathsf{C}_{\mathbb{N}}' <_{\mathrm{W}} \mathsf{K}_{\mathbb{N}}'' <_{\mathrm{W}} \mathsf{C}_{\mathbb{N}}'' <_{\mathrm{W}} \ldots$$

The so-called *compact choice problem* $\mathsf{K}_{\mathbb{N}}$ and the so-called *closed choice problem* $\mathsf{C}_{\mathbb{N}}$ on the natural numbers play an important role in Weihrauch complexity. As noted by the author and Rakotoniaina [BR17, §7] they can be seen as Weihrauch complexity analogs of the Kirby-Paris hierarchy

$$\mathsf{B}\Sigma_1^0 \leftarrow \mathsf{I}\Sigma_1^0 \leftarrow \mathsf{B}\Sigma_2^0 \leftarrow \mathsf{I}\Sigma_2^0 \leftarrow \mathsf{B}\Sigma_3^0 \leftarrow \mathsf{I}\Sigma_3^0 \leftarrow \ldots$$

of boundedness and induction principles as it is known from reverse mathematics.

Our main results on the computability-theoretic classification of the Weihrauch complexity of the Gödel problem and its variants is shown in in Figure 1. Additionally, we have results on the topological classification not shown here.

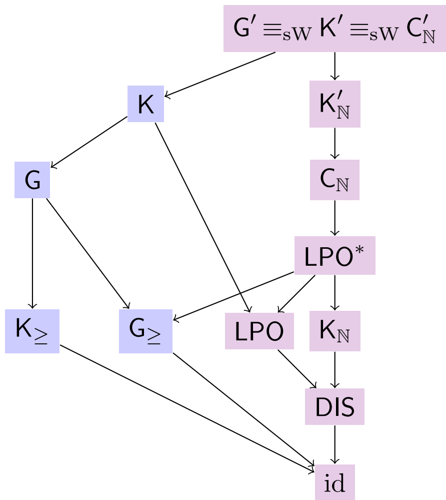

Figure 1. The Gödel problem in the Weihrauch lattice.

These results appeared in the conference article [Bra23b] and in a full version with all proofs in [Bra23a].

References

[BR17] Vasco Brattka and Tahina Rakotoniaina. On the uniform computational content of Ramsey's theorem. *Journal of Symbolic Logic*, 82(4):1278–1316, 2017.

[Bra23a] Vasco Brattka. On the complexity of computing Gödel numbers. arXiv 2302.04213, 2023.

[Bra23b] Vasco Brattka. On the complexity of learning programs. In Gianluca Della Vedova, Besik Dundua, Steffen Lempp, and Florin Manea, editors, *Unity of Logic and Computation*, volume 13967 of *Lecture Notes in Computer Science*, pages 166–177, Cham, 2023. Springer. 19th Conference on Computability in Europe.

[HR17] Mathieu Hoyrup and Cristóbal Rojas. On the information carried by programs about the objects they compute. *Theory of Computing Systems*, 61(4):1214–1236, 2017.

Faculty of Computer Science, Universität der Bundeswehr München, Germany and Dept. of Mathematics & Appl. Math. University of Cape Town, South Africa[1]

Email address: `Vasco.Brattka@cca-net.de`

[1]Vasco Brattka is supported by the National Research Foundation of South Africa.