On the complexity of learning programs

Vasco Brattka

Universität der Bundeswehr München, Germany
University of Cape Town, South Africa

Computability in Europe 2023
Batumi, Georgia
27 July 2023

Introduction

- Given a prefix of a sequence of numbers

$$3, 9, 15, 21, ...,$$

one can ask how the sequence continues?

- Provided the input sequence is total computable, the answer could be a Gödel number for it.
- This and similar questions have been intensively studied in algorithmic learning theory.
- Gold proved 1967 that one cannot even learn the Gödel number in the limit, in the situation above.
- We want to classify the Weihrauch complexity of the above problem.
- In this way we get a better understanding of the mixture of topological and computability-theoretic features that are involved in this problem.

Gödelization and Kolmogorov complexity

- Let $\varphi : \mathbb{N} \to \mathcal{P}$ be some standard Gödel numbering of the set $\mathcal{P}$ of partial computable functions.
- We call the following problem the Gödelization problem
$$\mathsf{G} :\subseteq \mathbb{N}^{\mathbb{N}} \rightrightarrows \mathbb{N}, p \mapsto \{i \in \mathbb{N} : \varphi_i = p\},$$
where $\mathrm{dom}(\mathsf{G})$ contains all total computable functions p.
- For our purposes the Kolmogorov complexity is the problem
$$\mathsf{K} :\subseteq \mathbb{N}^{\mathbb{N}} \to \mathbb{N}, p \mapsto \min \mathsf{G}(p),$$
with $\mathrm{dom}(\mathsf{K}) = \mathrm{dom}(\mathsf{G})$.
- Hoyrup and Rojas (2017) have coined the following slogan:
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.

Variants of the Gödelization problem

- We also look at the following variant of G:

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

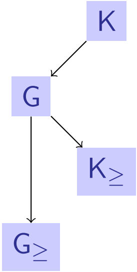
  where $\mathrm{dom}(\mathsf{G}) = \{(p, m) : K(p) \leq m\}$.

- And we study the following variant of K:

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

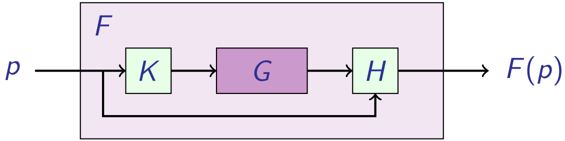
  with $\mathrm{dom}(\mathsf{K}_{\geq}) = \mathrm{dom}(\mathsf{G})$.

- These problems are related in the Weihrauch lattice as follows:



Weihrauch reducibility

Let $f :\subseteq X \rightrightarrows Y$ and $g :\subseteq Z \rightrightarrows W$ be two multi-valued functions.



- f is Weihrauch reducible to g, $f \le_{\mathrm{W}} g$, if there are computable $H, K :\subseteq \mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}$ such that $H\langle \mathrm{id}, GK\rangle \vdash f$ whenever $G \vdash g$.
- We write $f \le_{\mathrm{W}}^{*} g$ for the continuous version of Weihrauch reducibility, where H, K are chosen to be continuous.
- We write $f \le_{\mathrm{W}}^{p} g$ if H, K can be chosen to be computable relative to $p \in \mathbb{N}^{\mathbb{N}}$.
- $\equiv_{\mathrm{W}}$, $\equiv_{\mathrm{W}}^{*}$, and $\equiv_{\mathrm{W}}^{p}$ denote the corresponding equivalences.
- The distributive lattice induced by $\le_{\mathrm{W}}$ is usually referred to as Weihrauch lattice.

Typical problems in the Weihrauch lattice

- Limited principle of omniscience:
  $\mathsf{LPO} : \mathbb{N}^{\mathbb{N}} \to \{0,1\}, \mathsf{LPO}(p) = 1 :\iff p = \widehat{0}$
- Lesser limited principle of omniscience:
  $\mathsf{LLPO} :\subseteq \mathbb{N}^{\mathbb{N}} \rightrightarrows \{0,1\}, \mathsf{LLPO}\langle p_0, p_1\rangle := \{i \in \{0,1\} : p_i = \widehat{0}\}$,
  with $\mathrm{dom}(\mathsf{LLPO}) = \{\langle p_0, p_1\rangle \in \mathbb{N}^{\mathbb{N}} : \neg(p_0 \neq \widehat{0} \wedge p_1 \neq \widehat{0})\}$.
- Closed choice on $\mathbb{N}$ is
  $$\mathsf{C}_{\mathbb{N}} :\subseteq \mathbb{N}^{\mathbb{N}} \rightrightarrows \mathbb{N}, p \mapsto \{n \in \mathbb{N} : (\forall k)\ p(k) \neq n\},$$
  with $\mathrm{dom}(\mathsf{C}_{\mathbb{N}}) = \{p \in \mathbb{N}^{\mathbb{N}} : \mathrm{range}(p) \subsetneq \mathbb{N}\}$,
- Compact choice $\mathbb{N}$ is
  $$\mathsf{K}_{\mathbb{N}} :\subseteq \mathbb{N}^{\mathbb{N}} \times \mathbb{N} \rightrightarrows \mathbb{N}, (p, m) \mapsto \{n \leq m : (\forall k)\ p(k) \neq n\},$$
  with $\mathrm{dom}(\mathsf{K}_{\mathbb{N}}) = \{(p, m) \in \mathbb{N}^{\mathbb{N}} \times \mathbb{N} : \mathrm{range}(p) \subsetneq \{0, ..., m\}\}$.
- Weak Kőnig's lemma: $\mathsf{WKL} :\subseteq \mathsf{Tr} \rightrightarrows 2^{\mathbb{N}}, T \mapsto [T]$
- Limit: $\lim :\subseteq \mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}, \langle x_n\rangle \mapsto \lim_{n\to\infty} x_n$.

Borel complexity and Weihrauch complexity

The jump f' of a problem is a strengthening of f:

- a name of an input x for f' is a sequence (p_n) in $\mathbb{N}^{\mathbb{N}}$ that converge to a name $p \in \mathbb{N}^{\mathbb{N}}$ of an input in the sense of f.

Theorem (B. 2005, Pauly, de Brecht 2014 and Kihara 2015)

1. *f is computably Σ^0_{n+2}–measurable* $\iff$ $f \leq_{\mathrm{W}} \lim^{(n)}$.
2. *f is computably $(\Sigma^0_{n+2}, \Sigma^0_{n+2})$–measurable* $\iff$ $f \leq_{\mathrm{W}} \mathsf{C}_{\mathbb{N}}^{(n)}$.

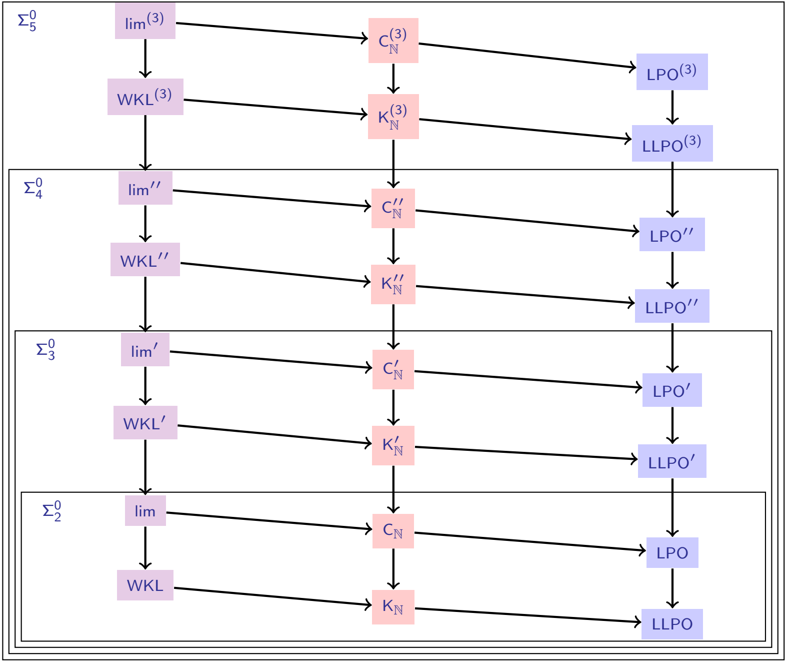
- Weihrauch complexity refines Borel complexity very much in the same way as many-one complexity refines arithmetical complexity.
- B. and Rakotoniaina (2017) have shown that

  $$\mathsf{K}_{\mathbb{N}} \leq_{\mathrm{W}} \mathsf{C}_{\mathbb{N}} \leq_{\mathrm{W}} \mathsf{K}'_{\mathbb{N}} \leq_{\mathrm{W}} \mathsf{C}'_{\mathbb{N}} \leq_{\mathrm{W}} ...$$

  and concluded that this is the proper Weihrauch analogue of the Paris-Harrington hierarchy of induction and boundedness problems

  $$\mathsf{B}\Sigma^0_1 \leftarrow \mathsf{I}\Sigma^0_1 \leftarrow \mathsf{B}\Sigma^0_2 \leftarrow \mathsf{I}\Sigma^0_2 \leftarrow ...$$

Basic skeleton of Weihrauch complexity



Motivation for closed and compact choice as benchmarks

Recall that the first-order part of a problem f can be defined by

$$^1f := \max_{\leq_W}\{g :\subseteq \mathbb{N}^\mathbb{N} \rightrightarrows \mathbb{N} : g \leq_W f\}.$$

It was introduced by Dzhafarov, Solomon, and Yokoyama (2019).

Theorem (Valenti 2021, Soldà and Valenti 2023)

1. $^1(\lim^{(n)}) \equiv_{sW} \mathsf{C}_\mathbb{N}^{(n)}$, *in particular* $^1\lim \equiv_{sW} \mathsf{C}_\mathbb{N}$,
2. $^1(\mathsf{WKL}^{(n)}) \equiv_{sW} \mathsf{K}_\mathbb{N}^{(n)}$, *in particular* $^1\mathsf{WKL} \equiv_{sW} \mathsf{K}_\mathbb{N}$.

By a result of Westrick (2021) the diamond can be characterized by

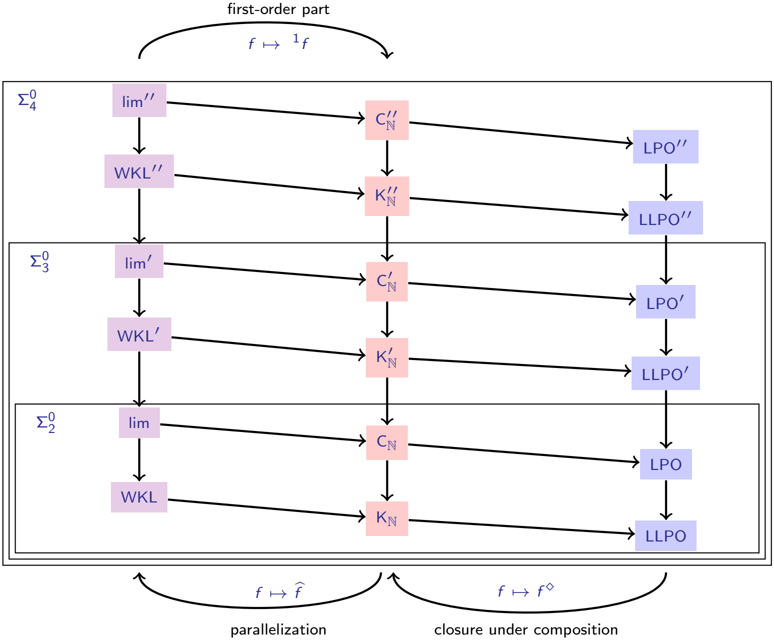
$$f^\diamond := \max_{\leq_W}\{g :\subseteq \mathbb{N}^\mathbb{N} \rightrightarrows \mathbb{N}^\mathbb{N} : f \leq_W g \star g \leq_W g\}.$$

It was introduced by Neumann and Pauly (2018).

Proposition

- $\mathsf{LPO}^\diamond \equiv_W \mathsf{C}_\mathbb{N}$ *(Neumann and Pauly 2018)*
- $\mathsf{LLPO}^\diamond \equiv_W \mathsf{K}_\mathbb{N}$ *(Soldà and Valenti 2023)*

Basic skeleton of Weihrauch complexity



Reverse mathematics and computability classes

Weihrauch degree	Reverse mathematics axioms
$C_{\mathbb{N}^\mathbb{N}}$	ATR_0
$\lim^\diamond$	ACA_0
WKL	WKL_0^*
$C_\mathbb{N}^{(n)}$	$I\Sigma_{n+1}^0$
$K_\mathbb{N}^{(n)}$	$B\Sigma_{n+1}^0$
id	RCA_0^*

Theorem (B., de Brecht and Pauly 2012)

1. *f is limit computable* $\iff$ $f \leq_W \lim$.
2. *f is finite mind change computable* $\iff$ $f \leq_W C_\mathbb{N}$.
3. *f is non-deterministically computable* $\iff$ $f \leq_W WKL$.

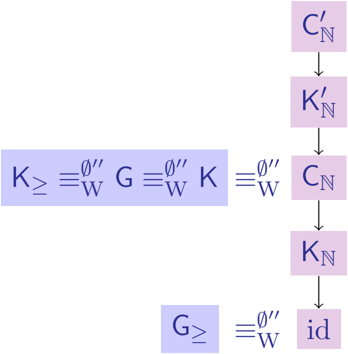
- Gold's result can be translated into $G \not\leq_W C_\mathbb{N}$.
- We will use the problems $K_\mathbb{N}$ and $C_\mathbb{N}$ as a benchmark to classify the Gödel problem.

The topological situation

$$\begin{array}{ccc}
 & & \mathsf{C}'_{\mathbb{N}} \\
 & & \downarrow \\
 & & \mathsf{K}'_{\mathbb{N}} \\
 & & \downarrow \\
\mathsf{K}_{\geq} \equiv^{*}_{\mathrm{W}} \mathsf{G} \equiv^{*}_{\mathrm{W}} \mathsf{K} & \equiv^{*}_{\mathrm{W}} & \mathsf{C}_{\mathbb{N}} \\
 & & \downarrow \\
 & & \mathsf{K}_{\mathbb{N}} \\
 & & \downarrow \\
\mathsf{G}_{\geq} & \equiv^{*}_{\mathrm{W}} & \mathrm{id}
\end{array}$$

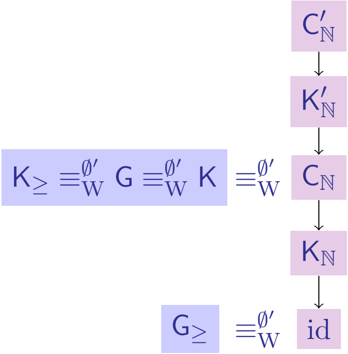
- The equivalence $\mathsf{K}_{\geq} \equiv^{*}_{\mathrm{W}} \mathsf{G}$ validates Hoyrup and Rojas slogan topologically.
- Which is the minimal oracle among $\emptyset, \emptyset', \emptyset'', \ldots$ that validates the picture above in place of $*$?

Optimal oracles



- The oracle $\emptyset''$ makes totality decidable and this yields easy proofs of the equivalences.
- Surprisingly, this can also be done with $\emptyset'$, but the proofs are slightly more difficult in this case.

Optimal oracles



- The oracle $\emptyset''$ makes totality decidable and this yields easy proofs of the equivalences.
- Surprisingly, this can also be done with $\emptyset'$, but the proofs are slightly more difficult in this case.

Upper bound with respect to the halting problem

Proposition

$\mathsf{K} \leq_{\mathrm{W}}^{\emptyset'} \mathsf{C}_{\mathbb{N}}$.

Proof.

- We go through all Gödel numbers $i = 0, 1, 2, \ldots$ one by one.
- For each i we check for each $n = 0, 1, 2, \ldots$ whether $n \in \mathrm{dom}(\varphi_i)$ (with the help of the halting problem) and whether $\varphi_i(n) = p(n)$.
- If so, then we write i to the output q and we move on to the next n.
- If one of these tests fails, then we move on to the next i.
- This procedure stops going to the next i when the smallest i with $\varphi_i = p$ is reached.
- Altogether, this gives a finite mind change computation for K.

□

Lower bound with respect to the halting problem

Proposition

$\mathsf{C}_{\mathbb{N}} \leq_{\mathrm{W}}^{\emptyset'} \mathsf{K}_{\geq}$.

Proof.

- We use a variant of the set of random natural numbers:
$$R := \{\langle k, n \rangle \in \mathbb{N} : \min\{i \in \mathbb{N} : \varphi_i(k) = n\} \geq n\}.$$
- For each k there are infinitely many n with $\langle k, n \rangle \in R$.
- R is co-c.e. and hence $R \leq_{\mathrm{T}} \emptyset'$.
- We use the boundedness problem $\mathsf{B} \equiv_{\mathrm{W}} \mathsf{C}_{\mathbb{N}}$, which is the problem: given a monotone increasing bounded sequence $p \in \mathbb{N}^{\mathbb{N}}$, find an upper bound $b \in \mathbb{N}$.
- We prove $\mathsf{B} \leq_{\mathrm{W}}^{R} \mathsf{K}_{\geq}$: inspecting the numbers $p(0), p(1), p(2), ...$ we construct $q(0), q(1), q(2), ...$ such that $b = \mathsf{K}(q)$ is an upper bound for p.
- This can be done such that q is eventually constant and hence actually computable.

□

Optimal oracles

$$\begin{array}{ccc}
 & & C'_{\mathbb{N}} \\
 & & \downarrow \\
 & & K'_{\mathbb{N}} \\
 & & \downarrow \\
K_{\geq} \equiv_{W}^{\emptyset'} G \equiv_{W}^{\emptyset'} K & \equiv_{W}^{\emptyset'} & C_{\mathbb{N}} \\
 & & \downarrow \\
 & & K_{\mathbb{N}} \\
 & & \downarrow \\
G_{\geq} & \equiv_{W}^{\emptyset'} & \mathrm{id}
\end{array}$$

- We have established the upper equivalences.
- We still need to prove $G_{\geq}$ is computable relative to the halting problem.

Computability with respect to the halting problem

Proposition

$\mathrm{G}_{\geq}$ *is computable with respect to the halting problem* $\emptyset'$.

Proof. We use a variant of the amalgamation technique.

- We consider the compatibility relation on $\mathcal{P}$:

$$f \approx g :\iff (\forall n \in \mathrm{dom}(f) \cap \mathrm{dom}(g))\ f(n) = g(n).$$

- $C := \{\langle i, j\rangle \in \mathbb{N} : \varphi_i \approx \varphi_j\}$ is co-c.e. and hence $C \leq_{\mathrm{T}} \emptyset'$.
- Let (p, m) be an input for $\mathrm{G}_{\geq}$, i.e., $\mathrm{K}(p) \leq m$.
- For $i \leq m$ that we consider the pockets:

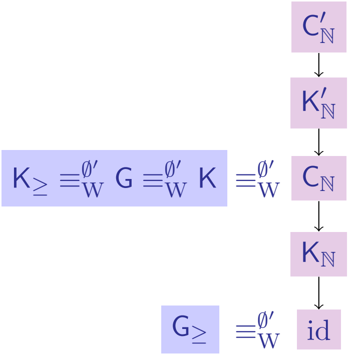
$$P_i := \{j \leq m : \varphi_i \approx \varphi_j\}$$

- P_i is called compatible, if $\varphi_{j_0} \approx \varphi_{j_1}$ holds for all $j_0, j_1 \in P_i$.
- Among $P_0, \ldots, P_m$ we remove all incompatible pockets and all double occurrences of the same pocket.
- This yields a list of $P_{i_0}, \ldots, P_{i_k}$ of pairwise different pockets, which are all compatible by themselves.

Computability with respect to the halting problem

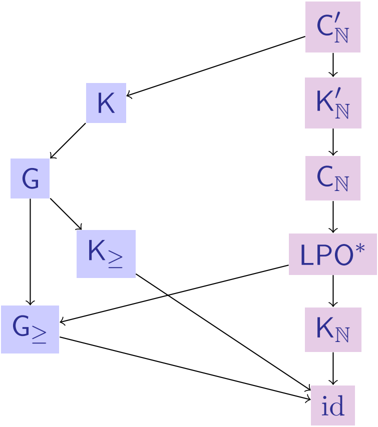
- No pocket in our list is a subset of another pocket.
- Among the pockets $P_{i_0}, \ldots, P_{i_k}$ in our list
  1. exactly one contains at least one code j with $\varphi_j = p$ and all codes j in this pocket satisfy $\varphi_j \approx p$,
  2. all other pockets contain at least one j with $\varphi_j \not\approx p$.
- P_i is called compatible with p, if $p \approx \varphi_j$ for all $j \in P_i$.
- 1. and 2. guarantee that there is exactly one pocket P_i among the $P_{i_0}, \ldots, P_{i_k}$ that is compatible with p and contains a Gödel number of p.
- A prefix of p is sufficient to identify P_i as we just need to find an incompatible member in all the other pockets.
- From the index i we can compute a Gödel number $r(i)$ of p: for each input $n \in \mathbb{N}$ we search for some $j \in P_i$ such that $n \in \mathrm{dom}(\varphi_j)$ and we produce $\varphi_j(n)$ as result.
- Hence, $r(i) \in \mathsf{G}_{\geq}\langle p, m\rangle$. (We note that $r(i) \leq m$ is not required and might not hold.) □

Optimal oracles



- We now want to study the situation in the computable case.
- We know $\mathsf{G} \not\leq_{\mathrm{W}} \mathsf{C}_{\mathbb{N}}$ by Gold (1967) and $\mathsf{G}_{\geq} \leq_{\mathrm{W}} \mathsf{C}_{\mathbb{N}}$ by Freivald and Wiehagen (1979).

The computability-theoretic situation



The computability-theoretic situation

- $\mathsf{K} \leq_{\mathrm{W}} \mathsf{C}_{\mathbb{N}}'$ can be proved observing that $\mathsf{C}_{\mathbb{N}}' \equiv_{\mathrm{W}} \liminf_{\mathbb{N}}$. We just write all Gödel numbers i onto the output that match the input for longer and longer prefixes of the input p. The least cluster point is the smallest Gödel number of p.
- $\mathsf{K}_{\geq} \not\leq_{\mathrm{W}} \mathsf{K}_{\mathbb{N}}'$ can be proved by a finite extension construction using that $\mathsf{K}_{\mathbb{N}}' \equiv_{\mathrm{W}} \mathsf{BWT}_{\mathbb{N}}$ (the Bolzano-Weierstraß theorem on $\mathbb{N}$).
- Hence the classification of $\mathsf{K}_{\geq} \leq_{\mathrm{W}} \mathsf{G} \leq_{\mathrm{W}} \mathsf{K}$ is optimal with respect to our benchmark problems.
- $\mathsf{G}_{\geq} \leq_{\mathrm{W}} \mathsf{LPO}^*$ can be proved with the amalgamation technique.
- $\mathsf{G}_{\geq} \not\leq_{\mathrm{W}} \mathsf{K}_{\mathbb{N}}$ can be proved with a finite extension construction.
- $\mathsf{G}_{\geq}$ is hence continuous, but not computable.
- The problems $\mathsf{G}_{\geq}, \mathsf{K}_{\geq}, \mathsf{G}$ and K can all be separated from each other with respect to $\leq_{\mathrm{W}}$.

Closure properties of Gödelization

- By $\widehat{\mathsf{G}}$ we denote the parallelization of G
- By $\mathsf{G} \star \mathsf{G}$ we denote the compositional product of G by itself
- By G^* we denote the finite parallelization of G
- By $f|_{\mathrm{c}}$ we denote the restriction to computable inputs of f

- $\widehat{\mathsf{G}}|_{\mathrm{c}} \equiv_{\mathrm{W}} \mathsf{G} <_{\mathrm{W}} \widehat{\mathsf{G}}$ (parallelization)
- $(\mathsf{G} \star \mathsf{G})|_{\mathrm{c}} \equiv_{\mathrm{W}} \mathsf{G}$ (compositional products)
- $\mathsf{G}^* \equiv_{\mathrm{W}} \mathsf{G}$ (finite parallelization)

Question

Does $\mathsf{G} \star \mathsf{G} \equiv_{\mathrm{W}} \mathsf{G}$ *hold?*

Lower bounds

Proposition

$\mathsf{DIS} \not\leq_{\mathrm{W}} \mathsf{G}$, *but* $\mathsf{LPO} \leq_{\mathrm{W}} \mathsf{K}$.

Proof. $\mathsf{DIS} \leq_{\mathrm{W}} \mathsf{G}$ would imply $\mathsf{NON} \leq_{\mathrm{W}} \widehat{\mathsf{G}}$, since $\widehat{\mathsf{DIS}} \equiv_{\mathrm{W}} \mathsf{NON}$. But since $\widehat{\mathsf{G}}|_{\mathrm{c}} \leq_{\mathrm{W}} \mathsf{G}$, this is impossible!
$\mathsf{LPO} \leq_{\mathrm{W}} \mathsf{K}$ is easy to see, as there is a specific smallest Gödel number i of the zero sequence $p \in \mathbb{N}^{\mathbb{N}}$. □

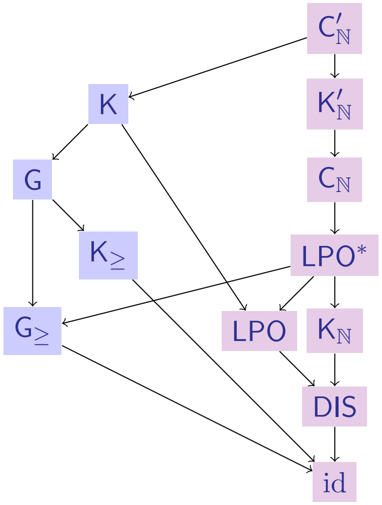
DIS is the weakest natural discontinuous problem with respect to topological Weihrauch reducibility (in $\mathsf{ZF}+\mathsf{DC}+\mathsf{AD}$). Hence, Gödelization G has no useful natural lower bounds (besides id)!

Corollary

G *is effectively discontinuous, but not computably so.*

This means $\mathsf{DIS} \leq_{\mathrm{W}}^{*} \mathsf{G}$, but $\mathsf{DIS} \not\leq_{\mathrm{W}} \mathsf{G}$.

The computability-theoretic situation



References

- V. Brattka, The Discontinuity Problem, *Journal of Symbolic Logic* (2022) (electronically published).
- V. Brattka and T. Rakotoniaina, On the Uniform Computational Content of Ramsey's Theorem, *Journal of Symbolic Logic* 82:4 (2017) 1278–1316.
- D. Dzhafarov, R. Solomon, and K. Yokoyama, On the first-order parts of Weihrauch degrees, in preparation, 2019.
- E.M. Gold, Language identification in the limit, *Information and Control* 10 (1967) 447–474
- M. Hoyrup and C. Rojas, On the information carried by programs about the objects they compute, *Theory of Computing Systems* 61:4 (2017) 1214–1236.
- E. Neumann and A. Pauly, A topological view on algebraic computation models, *Journal of Complexity* 44 (2018) 1–22
- G. Soldà and M. Valenti, Algebraic properties of the first-order part of a problem, *Annals of Pure and Applied Logic* 174 (2023) 103270.
- M. Valenti, A journey through computability, topology and analysis, University of Udine, PhD thesis, 2021.