

FOUNDATIONS OF ONLINE STRUCTURE THEORY II: THE OPERATOR APPROACH

ROD DOWNEY, ALEXANDER MELNIKOV, AND KENG MENG NG

ABSTRACT. We introduce a framework for online structure theory. Our approach generalises notions arising independently in several areas of computability theory and complexity theory. We suggest a unifying approach using operators where we allow the input to be a countable object of an arbitrary complexity. Thus we will begin a view that online algorithms can be viewed as a sub-area of computable analysis. We will give a new framework which (i) ties online algorithms with computable analysis, (ii) shows how to use modifications of notions from computable analysis, such as Weihrauch reducibility, to analyse *finite* but uniform combinatorics, (iii) show how to finitize reverse mathematics to suggest a fine structure of finite analogs of infinite combinatorial problems, and (iv) see how similar ideas can be amalgamated from areas such as EX-learning, computable analysis, distributed computing and the like. Conversely, we also get an enrichment of computable analysis from ideas from the analysis of classical online algorithms.

CONTENTS

1. Introduction	2
1.1. Our Goal	2
1.2. The Punctual Model	3
1.3. The Uniform/Operator Model	3
2. The main definition	6
2.1. Representation spaces	6
2.2. Representations	6
2.3. Online problems	6
2.4. Taking the completion of an online problem	7
2.5. The definition	7
2.6. The robustness lemma	8
2.7. Generalisations and refinements of the main definition.	10
2.8. General computable and efficient solutions	11
2.9. Multiple solutions	11
3. Oracle computation and uniformity	12
3.1. Graph oracles do not help	12
3.2. Interactions with punctual structure theory	14
4. Weihrauch reduction and online algorithms	15
4.1. Weihrauch reduction and incremental computation	16
4.2. Weihrauch reduction and online graph colouring	17

2010 *Mathematics Subject Classification.* Primary 03D45, 03C57. Secondary 03D75, 03D80.

The first two authors were partially supported by Marsden Fund of New Zealand. The third author is partially supported by the grant MOE2015-T2-2-055 and RG131/17.

5. Δ_2^0 processes, finite reverse mathematics, and Weihrauch reduction	23
6. Real functions.	25
References	27

1. INTRODUCTION

1.1. **Our Goal.** Imagine you are tasked with putting objects of differing sizes into bins of a fixed size. Your goal is to minimize the number of bins you need. This is the famous BIN PACKING problem which we know is NP complete (see Karp [38]). But imagine that we change the rules and you are only given the objects one at a time and you must choose which bin to put the object into before being given the next object. You are in an *online* situation and this is the ONLINE BIN PACKING problem. The “first fit” method is well-known to give a 2-approximation algorithm for this problem (definitions given in detail below). Alternatively imagine you are a scheduler, and your goal is to schedule requests within a computer for memory allocation amongst users. Again you are in an online situation, but here you might want to change the order of allocation depending on priorities of the requests. Or from algorithmic randomness, you have a (computable) KC-set of requests of the form $(2^{-n_i}, \sigma_i)$ with $\sum_{i=1}^{\infty} 2^{-n_i} \leq 1$, and need to build a prefix-free Turing machine M with strings τ_i such that $|\tau_i| = n_i$ and $M(\tau_i) = \sigma_i$. Then the proof from e.g. Downey and Hirschfeldt [17], Theorem 3.6.1 is online in the sense that for each request at step i we generate the string τ_i .

Thus an online algorithm is one which acts on a input which is given piece by piece in a serial fashion. In the case where the input is finite, Karp [39] suggested this as a sequence of “requests” r_1, r_2, \dots with the algorithm f specifying an action $f(r_1), f(r_1 r_2), \dots$. The natural model for this would be a database where a request would be an update. Note that this is quite distinct from the *offline* version where (in the finite case) the whole input is known in advance. It is important to realize that in practical online algorithms arising in computer science, the action needs to be specified before the next request is given. Occasionally this is varied with a *lookahead* or *delay* where typically we might get k further bits of input so r_1, \dots, r_{n+k} determines the next action, action number n is $f(r_1 r_2 \dots r_{n+k})$.

As we will see below, it is possible to generalize this further perhaps to have delay or lookahead $g(n)$ for action n with g some function of n (or even r_n) but this is not what happens in practice. Most natural examples, work from r_n to define $f(r_n)$ and indeed if r_n is a structure generated by $\{1, \dots, n\}$ of some kind, then $f(r_n)$ is some value $h(n)$.

A brief thought on this will reveal that there are potentially hundreds of situations where we are dealing with combinatorial algorithms for tasks where we only have partial evolving information about the the input data, or perhaps the data is so large that we cannot see it in total. This is the reason that there are so many algorithms for online tasks. Classical examples include insertion sort, perceptron, paging, job shop scheduling, ski rental, navigation with only local understanding, etc.; see [4]. On the other hand, there seems no general *theory* which we can use as a conceptual basis for the theory of (finite and infinite) online algorithms, and online

structures. Books in this area, such as Albers [4], all tend to be taxonomies of algorithms. Our goal is to give a theoretical basis for the theory of online algorithms and structures.

1.2. The Punctual Model. In [6] a related project was started aiming at providing a model-theoretical foundation to this theory. In that paper we focussed upon the intuition that online decisions in practice have *lack of delay*. That is, we need to pack the object into some bin immediately, before the next one is presented to us (in the BIN PACKING example). This led to a theory of online structures and algorithms we generally referred to as *punctual* structure theory.

What is the most general reasonable form of “punctual”? In [6] we gave several pages of analysis as to why there we chose to interpret punctuality as *primitive recursive*. That is, we chose primitive recursive as a *unifying abstraction* of the notion of lack of delay. We looked at models such as automata-based structure theory such as Khoussainov and Nerode [42], but noted that the approach is highly sensitive to presentations of the structures; for example treating the algorithms as generated by transducers yields a completely different theory to that obtained by treating functions as relations, as can be seen by comparing the approach of Khoussainov and Nerode, with that of Epstein et al. [22]. Also it would seem that although we can incorporate automatic processes in our theories, they are really no general enough for online algorithms in general. Similarly, polynomial time structures such as Cenzer and Remmel, Grigorieff, Alaev, and others [1, 2, 3, 13, 30] are rather presentation dependent. We also observe that many decision procedures in practice are primitive recursive without necessarily being polynomial time. Finally, primitive recursive has a nice Church-Turing thesis, in that it models computable processes without unbounded loops.

In [6], we imagine that our structures are given in a primitive recursive way and our goal is to build some desired function or object accordingly. This theory applies in many situations. In [6], we get a sub-field of computable structure theory, focussing on the punctuality of the demands of online-ness. This resulted in a surprisingly rich theory; e.g., [6, 18, 36, 37].

1.3. The Uniform/Operator Model. Imagine we need to build a colouring of a graph G which is given online. Thus, in the very simplest case, we would be given the graph $G = \lim_s G_s$, where G_s has s vertices. When the vertex s is introduced, we are also given at the same time precisely which vertices amongst $\{1, \dots, s-1\}$ has an edge with s (and this cannot change later). (This is the “request set” in Karp’s paper.) Our task is to colour s so that no two vertices which are connected have the same colour, before the opponent presents us with G_{s+1} .

Although in practice the task will be finite, since we have no idea how large the graph is, we can construe this as an infinite process. If we imagine this as an infinite process and we need to colour the whole of an infinite graph G . We can think of each possible version of G as being a path through an infinite tree of possibilities. Each node σ of length s of the tree will represent some graph G_σ with s vertices, and if $\sigma \prec \sigma'$ then G_σ is the subgraph of $G_{\sigma'}$ induced by vertices $\{1, \dots, s\}$. Note that there are only primitively recursively many non-isomorphic graphs with s vertices.

Then this view of an online algorithm differs from that given in [6] for the following core reason:

Although G can be viewed a path on an infinite primitive recursive tree of possibilities, there is no *a priori* reason that we should only consider a primitive recursive graph G . There are continuum many such paths and the online graph colouring problem can be considered for an infinite countable graph of any complexity.

The reader will quickly realise that the key point about online algorithms is one of *continuity* or *uniformity*. If we have a colouring of G_σ and we add a new vertex $s + 1$, the next G will be one of the possible extensions G_τ of G_σ with the vertex $s + 1$ added. For each such G_τ the colouring χ_{G_τ} must be compatible with the colouring χ_{G_σ} on G_σ .

1.3.1. *Computable Analysis.* The conclusion is that whilst online algorithms appear to be combinatorial algorithms on finite objects and possibly infinite ones, in fact they should be formulated as a branch of *computable analysis*. One of the goals of the present paper is to give such a formulation. We need to specify what kinds of spaces are of relevance and what kinds of operators correspond to online algorithms. We believe that this view will allow a discourse between the discrete and the continuous which could prove fruitful. Similar relationships between the continuous and the discrete have yielded powerful results such as in the Furstenberg view of Szemerédi's Theorem. Our unifying abstraction also means that computable analysis is shown to be important in finite combinatorics. We will also see that our abstraction means that we can relate the proof theory of finite combinatorics with classical proof theory, obtaining refinements of result for Reverse Mathematics, and we can relate the theory of incremental computation with ideas from computable analysis such as Weihrauch reducibility. Thus although this is not the most technically difficult paper, we see it as a conceptual advance showing that many ideas can be combined into a single unifying abstraction.

1.3.2. *Immediate actions.* Since we want the action to be immediate, following the abstraction of [6] and for the reasons above, we will also demand that the action works primitive recursively, or perhaps even running in polynomial time. Thus, for the example above, the most general online colouring algorithm must satisfy the following two features:

- χ_{G_τ} must agree with χ_{G_σ} for $\sigma \prec \tau$.
- The map $\tau \rightarrow \chi_{G_\tau}$ must be (minimally) *primitive recursive*.

The former clause corresponds to continuity of the solution, and the latter reflects that the action should be without delay as abstracted by primitive recursion. These will all soon be made precise and general using representations (i.e., naming systems) for online problems. In fact, there are at least three possible interpretations of the second clause above. For example, should we allow *lookahead* in the computation of f ? In Lemma 2.5 we will prove that all three potential definitions are equivalent up to a primitive recursive change of notation, and therefore our definition is *robust*.

We will also see that the earlier approach to online algorithms by Kierstead, Trotter et al. [44, 50, 45] and Borodin and El-Yaniv [8] can be viewed as a sub-recursive relativisation of our definition. According to this earlier approach, the map $\tau \rightarrow \chi_{G_\tau}$ just needs to be total and does not even have to be computable. So we see that primitive recursion is not that general when compared to some other definitions in the literature.

1.3.3. *Why primitive recursion?* In [6] and above we argued that primitive recursion captures the lack of delay in computable algebra. But it may seem at first glance that every reasonable online algorithm should be polynomial time. However, it is well-known that in practice some useful algorithms are (provably) not polynomial-time; yet they seem to perform well enough on most inputs. It is therefore not necessarily clear whether polynomial-time is the right abstraction for efficiency in the online situation. So we want our function to be in a nice complexity class but we are not yet sure what exactly this class should be. It makes sense to develop as much structural “punctual” theory as possible and then see how much of it is preserved when we restrict ourselves to some narrow complexity class. And if something fails, we will have a better idea what goes wrong in the worst possible scenario; e.g., we will compare the positive Theorem 6.3 and the analogous negative result in polynomial-time analysis [47]. Perhaps, it will help to define generic-case online algorithms in the future.

Also, as we have already mentioned above, the notion of a polynomial-time online procedure would be too notationally dependent to be unambiguous. One of the main advantages of our general primitive recursive approach is that we still can prove a number of notation-independent results; e.g., the “robustness” Lemma 2.5 and the above-mentioned Theorem 6.3. Such results focus more on the effects related to online-ness and less on the pathologies related to a specific choice of representation. On the other hand, we will also prove several results (e.g., Theorem 3.6) which show that sometimes *all* pathologies come from presentation because the only notation-independent online solutions are the trivial ones. Such results of the second kind will typically hold for polynomial-time or exponential (etc.) algorithms too, and via essentially the same proof.

1.3.4. *Computable analysis and uniformity.* As mentioned above, we realise that this material also has a connection with *computable* and *feasible analysis*, and also with the complexity theory for operators in analysis along the lines of Kawamura and Cook [40], Melhorn [51], Ko and Friedman [26], and others. We will also note connections with reverse mathematics, computational learning theory, and even algorithmic randomness. We will also see that, in this setting, the finiteness of the objects being given is not an essential restriction. In the online case, finite objects are only revealed one bit at a time, and for all intents and purposes, we may as well treat all inputs as arbitrarily large finite structures. Working with arbitrarily large finite structures and infinite structures are indeed the same for our setting. This allows for example, for a formal approach in which one can study finite combinatorics in reverse mathematics.

As a final remark, we mention that we see this work as an extension of [6] in the following way.

In [6] we considered online computation of primitive recursive structures, with primitive recursive functions. This is akin to the Turing-Markov [60] view of computable analysis as effective processes on the countable field of *computable* reals. The Gregorczyk-Kleene [29] view, called type II computability, which views computable analysis as effective operators acting on the continuum of *all* reals. It is also akin to the bifurcation between computable structure theory and *uniform* computable structure theory.

2. THE MAIN DEFINITION

In the following sections, we will work up to the main definition. Remember that we want to simultaneously generalize online algorithms on finite and infinite structures, and in a general setting where the domains might be any kind of structure. Because of this we will need to tour through *representations* (computational ways of naming infinite objects), and carefully argue why choices, such as that of primitive recursion, are made.

2.1. Representation spaces. A class \mathcal{C} of relational structures is called *inductive* if $A \in \mathcal{C}$ implies A has a *filtration* $A = \cup_s A_s$ where each A_n is finite, has universe $\{1, \dots, n\}$, and for all $n' > n$ the substructure induced by $\{1, \dots, n\}$ in $A_{n'}$ is A_n . More generally, for a fixed computable function g , we say that \mathcal{C} is *g -inductive* if it has a g -filtration meaning that each A_n has universe $\{1, \dots, g(n)\}$. Here we will sometimes write $O(h(n))$ -inductive for the case where g is $O(h)$. Our language will typically be finite and relational¹.

We refer to the substructure of A based on $\{1, \dots, n\}$ the substructure of *height* $h(n) = n$. In the example discussed above, the height n structures are the graphs with n vertices. Another example is considered by Khousainov [41] with a height function in his work on random infinite structures. Natural online structures have natural height functions.

By abusing notation, we will let $\mathcal{C}^{<\omega}$ denote the class of finite substructures of \mathcal{C} . There is also the natural induced topology. For example, in the graph case this would be compact and have the totally disconnected topology with basic open sets being the extensions of graphs of height n .

2.2. Representations. A *representation* of an inductive class \mathcal{C} of structures is a computable surjective function which acts faithfully (and continuously) in the sense that $\delta(\sigma) = C_n$ for $|\sigma| = n$ and $h(C_n) = n$, and if $\sigma \preceq \tau$ then $\delta(\sigma)$ is an induced substructure of $F(\tau)$. We can also extend this in the natural way to g -filtrations. Thus such a δ induces a map $\bar{\delta}$ from $\omega^\omega \rightarrow \mathcal{C}$, namely $\lim\{\delta\sigma \mid \sigma \prec x\}$. We will call $x \in \omega^\omega$ a *name* for $C \in \mathcal{C}$ if $\bar{\delta}(x) = C$.

For the time being, we will regard $\bar{\delta}$ as being injective. When it is possible, we will replace $\omega^{<\omega}$ with $2^{<\omega}$. We will consider functions $f : \mathcal{C}_1 \rightarrow \mathcal{C}_2$ and these are represented by functions F acting on representations Q_i of \mathcal{C}_i , so taking finite strings to finite strings.

We emphasise that the function F is acting on *strings* which are *finite objects*. These represent, e.g., graphs. It is the continuity of the action will induce a map \bar{F} which is the completion of the finite maps. If x is a member of that completion; such as an infinite graph represented by a path through the relevant representing tree, then $x = \lim_n \delta^{-1}(G_n)$. Where the representation is obvious, we will suppress the explicit mention of this in the below to encourage readability.

2.3. Online problems. Although our objects of study are not strings, we will implicitly identify them with their representations, in accordance with the previous subsection. In particular, if the representation space is compact then our objects can be identified with strings over a finite alphabet.

¹Richard Shore observed that the punctual case focusses attention upon functions and functional languages, whereas the operator approach seems to tie itself to relational ones. We need some care if the language has function symbols, see [35].

Intuitively, to solve a problem we need to find a function f which, on input i , chooses an admissible solution from the finite set $s(i)$ of “correct” solutions.

Definition 2.1. An online problem is a triple (I, S, s) , where I is the space of inputs (i.e. the filtration) viewed as finite strings in a finite or infinite computable alphabet, S is the space of outputs viewed as finite strings in (perhaps, some other) alphabet, and $s : I \rightrightarrows S^{<\omega}$ is a (multi-)function which maps each $\sigma \in I$ to the set $s(\sigma)$ of admissible solutions of σ in S .

Note that the multi-valued function s does not have to be computable in general. For instance, for a colouring problem I will be codes for finite graphs and S for finite coloured graphs. Then $s(\sigma)$ will correspond to the collection of all admissible colourings; e.g., such that adjacent vertices are distinctly coloured. These colourings will form the space of admissible solutions.

Most natural problems from finite structures will obey the following convention, which we will consider in this section. Only in Section 6 will be consider more general cases.

Convention 2.2. Unless explicitly mentioned, I and S are compact with a primitive recursive modulus of compactness; i.e., it is primitively recursively branching when viewed as a tree of strings. Thus, there is a natural primitive recursive way to transform I into 2^ω (typically not height preserving).

2.4. Taking the completion of an online problem. A solution f to an online problem (I, S, s) induces a solution for the (topological) completion of the initial problem (I, S, s) , in the sense that f can be uniquely extended to a functional $\bar{f} : [I] \rightarrow [S]$. Here $[I]$ consists of infinite strings ξ such that for every i , $\xi \upharpoonright i \in I$, and similarly for $[S]$.

In general, in Definition 2.1 we may also require \bar{f} to satisfy some global property which cannot be always captured by s from Definition 2.1. For example, in Section 3 a solution must be an isomorphism between two presentations of the same infinite graph. In general, even if at every stage $f(\sigma)$ may be extendable to some isomorphism, the map associated with \bar{f} may fail to be surjective in the limit. Also, in another example in Section 3 we will require our solution to work only if the input is a presentation of some fixed infinite graph, which is also a property of \bar{f} rather than of any finite approximation to it. In particular, in this case admissibility of \bar{f} cannot be captured by s in Definition 2.1; at least not in general.

Convention 2.3. We will refer to such properties of \bar{f} as *global* and will not incorporate them into Definition 2.1.

2.5. The definition. Recall Conventions 2.2 and 2.3.

Definition 2.4. A punctual solution to (a representation of) an online problem (I, S, s) is a computable function $f : I \rightarrow S$ with the properties:

- (O1) $f(\sigma) \in s(\sigma)$ for every $\sigma \in I$;
- (O2) If $\sigma \prec \tau$ then $f(\sigma) \preceq f(\tau)$.
- (O3) f is primitive recursive.

Condition (O1) says that the output of f is an admissible solution. In (O2) we ask for is that each increment of the input yields an increment in the output, in the sense that $f(\sigma)$ must be a solution to σ .

The reader should note that (O3) is somewhat ambiguous as stated because it may be interpreted in at least two different ways, namely f could be primitive recursive either as a function or as a functional:

- (O3)' the computation of $f(\sigma)$ is based solely on σ , or
- (O3)'' the computation of σ may ask for an extension τ of σ before it halts.

Indeed, for online computations it would be natural to demand that we have a primitive recursive timestamp function g and to compute $f(\sigma)$ we would look at σ' of length $g(|\sigma|)$ extending σ . In practical computations lookahead will typically be $g(|\sigma|) = |\sigma| + k$ for some constant k . On the other hand, for a recursion theorist it would be more natural to consider *Turing functionals* acting on the representation spaces and demand that they are primitive recursive. By that we mean adding the characteristic function for the infinite string in the completion of the problem (§2.4) to the primitive recursive scheme of f ; to be clarified in Subsection 2.6. These two general definitions of lookahead (via timestamp and via oracle) are not equivalent when, say, $I \cong \omega^{<\omega}$. Thus, we have three natural versions of (O3) which are furthermore provably not equivalent in general.

Luckily, in the next subsection we will prove that, under Convention 2.2, these three versions of the main definition are equivalent up to a primitive recursive change of notation, and therefore Definition 2.4 is robust.

2.6. The robustness lemma. As we mentioned above, there are two natural ways of interpreting what it means for f in (O3) to be primitive recursive with a lookahead. We give more details.

In the first definition, we require that it is a Turing functional that possesses a primitive recursive time-function t which, on every input σ outputs the number of steps which f takes to compute $f(\sigma)$. In particular, $t(\sigma)$ bounds the use of the operator, that is, the length of τ extending σ which may be used in the computation of $f(\sigma)$. The length of the output $f(\sigma)$ will also be bounded by $t(\sigma)$.

The seemingly more general definition of a primitive recursive functional says that, for each infinite path x through the space of inputs, f is primitive recursive relative to $x = \lim_s \{\sigma \mid \sigma \prec x\}$. The latter can be formally defined by adding the characteristic function for x to the primitive recursive schema, and hence would potentially entail that $f(\sigma)$ could be arbitrarily long for various extensions of σ .

These two notions are equivalent in our framework. (Recall Convention 2.2.)

Lemma 2.5. *For a primitively recursively branching I , a Turing functional $f : I \rightarrow S$ possesses a primitive recursive time-function iff f is a primitive recursive functional. Moreover, if f possesses a primitive recursive time-function, then there is an equivalent online problem (I', S', s') with a primitive recursive solution without lookahead.*

Proof. In a different terminology the proof will appear in [35]. A similar formal argument can be found in the appendix of [6].

Suppose the Turing functional f possesses a primitive recursive time function t . Using t as a universal bound on all the searches which may occur in a computation with any oracle x extending σ , we can transform the general recursive scheme (augmented with the characteristic function χ_x for x) into a primitive recursive scheme augmented with χ_x . This implication holds in general, i.e., without any extra assumption on I .

Now, assuming I is primitively recursively branching, suppose f is a primitive recursive functional with functional oracle g in the most general relativised sense.

For the base of induction consider the following cases: $\Phi^g = o$, $\Phi^g = s$, $\Phi^g = I_m^n$, and $\Phi^g = g$. The first three cases are evident since they do not refer to g , while in the case when $\Phi^g = g$ take $t = b$, where b is the primitive recursive branching of I . Take $t(x) = \sum_{i \leq x} b(i)$ to make t monotonically increasing in its input.

The inductive step splits into two different cases depending on whether the last iteration is composition or an instance of primitive recursion.

Suppose it is composition,

$$\Psi^g(\bar{x}) = \Phi^g(\Theta_1^g(\bar{x}), \dots, \Theta_m^g(\bar{x})),$$

where $\Psi, \Theta_1, \dots, \Theta_m$ there are primitive recursive operators with primitive recursive time bounds t_0, t_1, \dots, t_m . As usual, we identify a tuple $\bar{x} = \langle x_1, \dots, x_m \rangle$ with its primitive recursive code. The time functions t_0, t_1, \dots, t_m can be assumed monotonically increasing in their inputs.

Define a primitive recursive time bound for Ψ by the rule

$$t(\bar{x}) = t_0(\langle t_1(\bar{x}), \dots, t_m(\bar{x}) \rangle) + \sum_i t_i(\bar{x}),$$

which can be rewritten into a primitive recursive schema using the standard techniques. Now suppose that Ψ is defined using an instance of primitive recursion, more specifically

$$\Psi^g(\bar{x}, 0) = \Theta^g(\bar{x});$$

$$\Psi^g(\bar{x}, y + 1) = \Phi^g(\bar{x}, y, \Psi^g(\bar{x}, y)),$$

where Φ and Θ are primitive recursive operators which have corresponding primitive recursive time functions t_0 and t_1 . Define t by the rule

$$t(\bar{x}, 0) = t_0(\bar{x});$$

$$t(\bar{x}, y + 1) = t_1(\bar{x}, y, t(\bar{x}, y)) + t(\bar{x}, y);$$

assuming that t is monotonically increasing in its input this gives the desired upper bound.

The last part of the lemma, recall that the tree I is primitively recursively branching. Thus, we can inductively form a new tree I' whose level n nodes are in a (primitive recursive) 1-1 correspondence with the nodes at level $t(n) = \max\{t(\sigma) : |\sigma| = n\}$ in I . Then the algorithm f on I works on I' as a strict algorithm. \square

Of course, if f is primitive recursive functional without lookahead then f can be viewed as simply a primitive recursive function mapping finite strings to finite strings. More formally, we have:

Fact 2.6. *Suppose $\mathcal{P} = (I, S, s)$ is an online problem. Then the following are equivalent:*

1. \mathcal{P} has a solution witnessed by a primitive recursive function f .
2. The completion of \mathcal{P} has a solution witnessed by a primitive recursive operator \bar{f} without lookahead.

2.7. Generalisations and refinements of the main definition. It is important to understand that primitive recursion smoothes many difficulties related to notation. In particular, the robustness lemma from the previous section will typically fail for polynomial-time operators. Thus, different interpretations of (O3) in Definition 2.4 will potentially lead to different refinements of the main definition to more narrow complexity classes. On the other hand, different versions of relativisation (such as general Turing and sub-recursive) will lead to potentially non-equivalent generalisations of the main definition. We will not develop these topics in too much detail, but some notions and notation introduced in this subsection will be important in the later sections.

2.7.1. Strict solutions, obT operators, and totality. We may want to stick with a given notation (I, S, s) because it may be either inconvenient to do so or perhaps computationally too hard. If the space I is not primitively recursively branching or not even compact, then Lemma 2.5 no longer holds. Thus, in this case the most general version of Definition 2.4 becomes ambiguous. In contrast, Fact 2.6 does not rely on compactness of I , let alone its primitive recursive branching, and therefore the stronger strict version of Definition 2.4 still makes sense even for non-compact I . Thus, the situation described in Fact 2.6 deserves a special attention.

Definition 2.7. In the simpler situation that no lookahead is allowed in Definition 2.4 we will call f a *strict* punctual solution.

By Fact 2.6, this situation can be considered an analog of a classical ibT-reduction, but acting on compact spaces with primitive recursive branchings with the branches of level n being the structures of height n , instead of 2^ω . Classically, ibT refers to an oracle procedure $\Gamma^B = A$ with the use $\gamma(x) = x$ for all x , and here we are identifying sets with their characteristic functions as usual. ibT functionals and the induced reduction have been studied quite intensively [15, 5, 17, 19, 59] and even used in (classical) differential geometry [15, 53].

Definition 2.8. For a fixed filtration representing I we will call such a procedure Γ induced by a strict online solution an *obT* (online bounded Turing) reduction.

Of course, the classical ibT reduction is usually viewed as working on 2^ω . If our space does not have primitive recursive branching then we no longer can transform it effectively into a copy of 2^ω . But as mentioned earlier, we see this aspect as a feature of the model, and not a flaw. One should expect online-ness to be generally representation dependent, at least to some extent.

Although there are notions of a polynomial-time functional in the literature [47], Definition 2.7 is much more convenient if we want to define what it means for a punctual solution to be polynomial time.

We can also use Definition 2.7 to give an explicit connection of our definition with the above-mentioned approach in [44, 8] which relies on total (not necessarily computable) functions. As has been observed in [35], a total function can be viewed as a function primitive recursive relative to some oracle. More formally, we have:

Fact 2.9. *For an online problem, the following are equivalent:*

- (1) *The problem has a total strict solution f (in the sense of [44, 8]);*
- (2) *The problem has a strict solution primitive recursive relative to some oracle (in the sense of [35]).*

In computability theory many arguments tend to be uniform enough to be relativisable to any oracle. The relativisation phenomenon partially explains why the seemingly crude approach via totality [44, 8] often captures some features of online computation.

2.8. General computable and efficient solutions. *Many* of our results are valid for computable online solutions, and in fact for any total solutions. For example, certainly a result showing that *no* computable solution is possible is very strong. For example, the proof that online colouring forests from Gasarch [28] requires $\Omega(\log n)$ many colours shows that no computable f is possible. We arrive at the following generalisation of the main definition.

Definition 2.10. A computable solution to (a representation of) an online problem (I, S, s) is a *Turing computable* function $f : I \rightarrow S$ with the properties (O1) and (O2) of Definition 2.4.

Should we use the definition above in place of our primitive recursive main definition? Turing computability is more general than primitive recursion. Some analogy of the robustness lemma (Lemma 2.5) will still hold. The key difference here is that it would hold for a completely different reason. A recursion theorist will be well-aware of how much unbounded search is abused in many such proofs. For example, we can use compactness of the representation space and *wait* for it to be covered by open sets. There will perhaps be no bound on how long we will have to wait, but from the point of view of Turing computability it will not make any difference. Primitive recursion seems just general enough for many structural results to hold, but often via a different, more subtle argument which takes into account punctuality of our procedure.

Nonetheless, there is one potential extra feature which is captured by this generalisation is that it also covers *partial* solutions and, potentially, *partial* representations. We will not make it overly formal and leave this to the reader. For example, in the case that the representations are partial, f would perhaps only need to work well on a valid input (cf. Convention 2.3). For example, in the case when ω^ω is representing Cauchy sequences, imagine we are seeing an online way to compute some continuous function. Then we would only need to produce a solution for those sequences which actually corresponded to convergent sequences. We could then require this solution be in some sense punctual when restricted to valid inputs.

We could on the other hand make the definition more efficient, for example:

Definition 2.11. A polynomial-time solution to (a representation of) an online problem (I, S, s) is a punctual strict solution which is furthermore polynomial-time.

The definition above is of course heavily notation-dependent. For the look-ahead case the situation becomes even more complex. Although there are definitions of a polynomial-time functional in the literature [47] they tend to be unconvincingly technical. Also, recall that the robustness lemma fails for polynomial-time simply because changing notation tends to be exponential time. Therefore the definition will also depend on which version of the main punctual definition we choose to make polynomial-time; recall there were three such versions.

2.9. Multiple solutions. Notice that in actual practice, we might also need a further generalisation of the above. Sometimes we might compute a (bounded)

collection of solutions at least one of which is correct at any stage and at height n . This occurs in, for example, using automata to compute minimization problems for graphs of bounded pathwidth (or k -interval graphs, see section 4.2) given the path decomposition. We will be computing a table of $f(k)$ many solutions at each level n . For example, for finding maximal clique you would have a collection of 2^k many possible solutions. However, it appears that a suitable choice of the space of outputs S can cover this seemingly more general case too.

3. ORACLE COMPUTATION AND UNIFORMITY

3.1. Graph oracles do not help. The main goal of this subsection is to show that a graph-oracle cannot significantly help in computing a function online. For that, we consider online functionals and online oracle computations.

Definition 3.1. We say that $f : 2^\omega \rightarrow 2^\omega$ is *online computable* if f has a representation $F : 2^{<\omega} \rightarrow 2^{<\omega}$, which is online computable in the sense above, so that for all $\alpha \in 2^\omega$, $F(\alpha \upharpoonright u(n)) = F(\alpha) \upharpoonright n$, where $F(\alpha) = \lim\{F(\sigma) \mid \sigma \prec \alpha\}$ and u is primitive recursive.

The space 2^ω can be replaced with a primitively recursively branching totally disconnected space. Identifying f with its representation F , we can unambiguously write this as $f(\alpha \upharpoonright u(n)) = f(\alpha) \upharpoonright n$, and (in view of Lemma 2.5) this should cause no problems in the case of primitively recursively branching spaces of strings. We may also allow more than one input in f .

Notation 3.2. It is natural to write $f^{\alpha \upharpoonright u(i)}(i)$ instead of $f(\alpha \upharpoonright u(i))$ and view α as an oracle. The output of $f^{\alpha \upharpoonright u(i)}(i)$ can also be interpreted as a natural number, when necessary.

Remark 3.3. There are obvious refinements of this. For example, it is natural to restrict ourselves to functionals f whose running time is a polynomial in the length of α . Also, having in mind some particularly nice primitive recursive function u , f is *u -online computable* if $f(\alpha \upharpoonright u(n)) = f(\alpha) \upharpoonright n$. An obvious case is when $u(n) = n + k$, which would be *online with delay k* . An illustration of this notion can be seen from Section 6, where we look at online real valued functions. We note that addition of reals is online computable with delay 2, meaning that to compute the sum of x and y to within 2^{-n} needs x and y with precision $2^{-(n+2)}$. Similar delay considerations come from other procedures in polynomial time analysis such as integration (see [47]). When $u(n) = n$ then the notions can be restated in terms of strict (ibT primitive recursive) functionals, while online with delay k corresponds to Lipschitz reducibility. Computable Lipschitz reducibility comes from algorithmic randomness ([17], Chapter 9) where it is shown that if f is online computable Lipschitz acting on 2^ω , then it preserves the Kolmogorov complexity of all sequences in the sense that for all n , $K(\alpha \upharpoonright n) \geq^+ K(f(\alpha) \upharpoonright n)$; that is $K(\alpha \upharpoonright n) \geq K(f(\alpha) \upharpoonright n) \pm O(1)$.

We will consider online functionals acting on algebraic or combinatorial structures, e.g., α could be viewed as a description of a finite segment of an infinite structure of some fixed finite relational signature, e.g., a graph. The extensions of $\alpha \upharpoonright n$ are the finitely many possible relational structures on $n + 1$ elements extending the structure described by $\alpha \upharpoonright n$. The intuition is that $f^{\alpha \upharpoonright u(i)}(i)$ is expected

to compute correctly only if α is an initial segment of a graph G . This is a global property; see Convention 2.3.

Definition 3.4. A function $h : \mathbb{N} \rightarrow \mathbb{N}$ is online computable from the isomorphism type of a structure G if there is an online f such that, whenever α is a description of G , $h(i) = f^\alpha(i)$.

In other words, h is allowed to use any presentation of some fixed G as its (online) oracle.

Example 3.5. To see how much extra computational power algebraic oracles can give, consider the following example. Let X be an arbitrary subset of \mathbb{N} , and define $A(X)$ to be an algebraic structure in the language of one unary function s , one unary predicate p , and one constant o , and which has the following isomorphism type. When restricted to s and o , it is just \mathbb{N} with $s(x) = x + 1$ and o interpreted as 0. Now define $p(x) \iff x \in X$. Given any presentation α of $A(X)$, we can decide X . So, in particular, computation from an isomorphism type is potentially as powerful as just the usual oracle computation.

In view of the example above, the reader will likely find the theorem below unexpected. Its proof is however not difficult; it can be viewed as a variation of an argument in Kalimullin, Melnikov, and Montalbán [35].

Theorem 3.6. *A function h is online computable from the isomorphism type of an infinite graph G if, and only if, h is primitive recursive.*

Remark 3.7. It will be clear from the proof below that the result has a natural polynomial time version. The exact definition of a polynomial time functional is a bit lengthy; see [26, 40]². We leave the polynomial time case to the reader.

Proof. By Ramsey's theorem, G either has an infinite clique or an infinite anti-clique; without loss of generality, suppose it is a clique. Since $g(i) = f^{\alpha \upharpoonright u(i)}(i)$, where α is *any* representation of G , we can assume that the first $u(i)$ bits of α describe a clique. Since the space of all presentations of G is primitively recursively branching, the use u is primitive recursive (see Lemma 2.5). Thus, the oracle can be completely suppressed and the trivial description of an infinite clique can be incorporated into a new procedure f_0 which does not use any oracle. On input i the procedure produces a string of length $u(i)$ which describes a finite clique, and then refers to this finite string (viewed as a partial function) whenever it needs to use the characteristic function of the oracle. This procedure is easily seen to be primitive recursive (as a function). \square

Informally, the result says that, from the perspective of online computation, graphs cannot code any non-trivial information into their isomorphism type; i.e., up to a change of their presentation. Both the theorem above and the main result in [18] imply that graphs are *not universal for punctual computability* – a notion which we will not formally define here (see [6]). See Kalimullin, Melnikov, and Montalbán [35] for a generalisation of Theorem 3.6 to structures in an arbitrary finite relational language.

²The point is that care is needed with which representations are allowed. Polynomial time functionals for $(0, 1)$ typically use the so-called signed digit representation, but even for \mathbb{R} there is some problem with the notion of the size of the input as discussed in, for instance, [40]. However, for any reasonable representation of graphs of size n this becomes relatively straightforward using, e.g. the standard matrix representation as in [27].

3.2. Interactions with punctual structure theory. In [6] we described the foundations of online structure theory. The main objects in this theory are infinite algebraic structures in which operations and relations are primitive recursive. As we argued in [6], there are natural strong connections of this new theory and the theory of polynomial-time algebraic structures (see also Alaev [1] and Alaev and Selivanov [3]) with applications to automatic structures [7]. Earlier we argued that this kind of punctual structure theory is akin to Turing-Markov computable analysis, in the the objects are given effectively. In this paper structures themselves do not have to be primitive recursive. However, the frameworks are closely related via, e.g., Theorem 3.9 below.

A *presentation* of a countably infinite algebraic structure in a finite language is an isomorphic copy of the structure upon the domain \mathbb{N} . For simplicity, we may assume that the structures in this section are all relational. In this case it becomes consistent with our framework; in particular, the space of all presentations I of a fixed structure in a finite relational language is primitively recursively branching.

Each such presentation $\alpha \in [I]$ can be viewed as an isomorphic copy of the structure upon the domain of \mathbb{N} . Some of these presentations will be *computable* in the sense that the relations on α will be computable predicates over \mathbb{N} . It is well-known that a structure may have non-computably isomorphic computable presentations. When we restrict ourselves to primitive recursive presentations and primitive recursive isomorphisms the situation becomes even more complex because the inverse of a primitive recursive function does not have to be primitive recursive. See [6] for a detailed exposition of the theory of punctually categorical structures.

The following notion is not restricted to primitive recursive presentations. A more general version of the definition below was first discussed briefly in [37] and then also mentioned in [36]. An even more general model-theoretic version of the definition can be found in [35].

Definition 3.8. *A structure G is strongly online categorical if there is an online strict operator f which, on input α and β arbitrary representations of G outputs an isomorphism from α onto β .*

In other words, there exists a primitive recursive functional $f^{\alpha;\beta}$ with both uses being the identity function, such that the associated function $h(i) = f^{\alpha \upharpoonright i; \beta \upharpoonright i}$ (whose output is interpreted as a natural number) induces an isomorphism from α onto β ; recall the latter two are isomorphic copies of G upon the domain \mathbb{N} . Equivalently, we could replace the functional by a primitive recursive function of three inputs σ, τ, i where $|\sigma| = |\tau| = i$ and finite strings are identified with their indices (under some fixed natural enumeration).

The theorem below can be viewed as a variation of another result of Kalimullin, Melnikov, and Montalbán [35] on punctual categoricity, but in our strongly online case the proof will be significantly simpler. Recall that a structure G is homogeneous if for any tuple \bar{x} in G and any pair of elements $y, z \in G$, we have that y is automorphic to z over \bar{x} .

Theorem 3.9. *A structure in a finite relational language is strongly online categorical if, and only if, it is homogeneous.*

Proof. Each homogeneous structure is trivially strongly online categorical. Now suppose G is strongly online categorical. Suppose the structure is not homogeneous, and let \bar{x} be shortest (of length n) such that for some z, y we have that z is not

in the same automorphism orbit as y over \bar{x} . Construct α and β as follows. First, copy \bar{x} into both and calculate the online isomorphism f from $\alpha \upharpoonright n$ to $\beta \upharpoonright n$. If we identify $\alpha \upharpoonright n$ and $\beta \upharpoonright n$ with \bar{x} , then f induces a permutation of $\beta \upharpoonright n$; by the choice of n any permutation of \bar{x} can be extended to an automorphism of the whole structure. Adjoin z to α and find a y' which plays the role of y over $\beta \upharpoonright n$ under any automorphism extending the permutation $\beta \upharpoonright n \leftrightarrow f(\alpha \upharpoonright n)$. Then necessarily $f(z) = f(y')$, because f has already shown its computation on the first n bits. However, by the choice of z and y' , f cannot be extended to an isomorphism no matter how we extend the presentations further. \square

Note that we used only totality of the strict functional in the proof. In the case when the language has functional symbols the theorem no longer holds. Of course, the notion of strongly online and of a presentation will have to be adjusted. But regardless, strong homogeneity will no longer capture the property (whatever it may be exactly).

Example 3.10. Consider the structure in the language of only one unary functional symbol s , and which consists entirely of disjoint 2-cycles. Here a 2-cycle is of course a component of the form $\{x, s(x)\}$ where $s(s(x)) = x$ and $x \neq s(x)$. According to any reasonable definition of (strong) online categoricity for functional structures, this structure has to be (strongly) online categorical. However, it is not homogeneous.

We leave open:

Question 1. Is it possible to find a reasonable algebraic description of (strongly) online categorical algebraic structures in an arbitrary finite language?

We suspect that such a description exists, and that the solution will likely boil down to setting the definitions right. If we replace strict with primitive recursive operators in Definition 3.8 we will obtain the more general notion of (uniform) online categoricity. With quite a bit of effort Theorem 3.9 can be extended [35] to this more general notion, and even beyond.

4. WEIHRAUCH REDUCTION AND ONLINE ALGORITHMS

Weihrauch reduction is one of the central notions in computable analysis. It was named by Brattka and Gherardi [11], before that it had been called recursive reducibility. Weihrauch reducibility can be viewed as a natural generalisation of computable Wadge reducibility [61]. Henceforth will use $f \leq_W g$ to denote Weihrauch reducibility. $f \leq_W g$ has the following intuition. We have some problem we wish to solve by computing an instance $f(x)$ some function f . To do this we produce another problem x' and solve $g(x')$ for g , and then convert $g(x')$ back to of $f(x)$. In more detail, we for functions f and g on ω^ω -represented spaces X and Y , $f \leq_W g$, is defined to mean that there are computable A and B on ω^ω , such that for *any* p_x , and *any* representation G of g ,

$$A(p_x, G(B(p_x)))$$

realizes f (i.e. is a name for $f(x)$). (This is defined here for single-valued functions, but does have a multy-valued version we won't need.) This should be thought of as follows for the archetypal case of a computable metric space. We computable metric space, we take a Cauchy sequence converging to x , use B to convert this

into a one converging to $B(x)$, and hence one converging to $g(B(X))$, and finally using the one converging to x and this one, to one converging to $A(x, g(B(x)))$.

The definition has a number of natural variations; some of these will be discussed below.

4.1. Weihrauch reduction and incremental computation. In this subsection we establish a formal connection between computable analysis and computer science. More specifically, we show that a version of Weihrauch reduction borrowed from computable analysis [62] is equivalent to incremental reduction between online problems suggested in Miltersen et al. [52].

We first state Weihrauch reductions in the online setting. Suppose \mathcal{P} , \mathcal{Q} are online problems.

Definition 4.1. *We say that \mathcal{P} is strongly Weihrauch reducible to \mathcal{Q} , written $\mathcal{P} \leq_{sW} \mathcal{Q}$, if there exist Turing functionals Φ and Ψ such that, whenever $\sigma \in I_{\mathcal{P}}$ is an instance of \mathcal{P} , $\Phi^\sigma = \tau \in I_{\mathcal{Q}}$ is an instance of \mathcal{Q} , and whenever $\rho \in s(\Phi^\sigma)$ is a solution to Φ^σ then $\theta = \Psi^\rho \in s(\sigma)$ is a solution to σ .*

Here the reduction is strong in the sense that there is a provably more general definition of (plain) Weihrauch reduction which will be given in due course. Note that, according to the definition above, all functionals involved are strict, but this condition can be relaxed giving a less tight reduction.

Notation 4.2. We write $\mathcal{P} \leq_{sW}^C \mathcal{Q}$ if both strict functionals (in our sense) Φ and Ψ in the definitions above belong to a complexity class C having sufficiently strong closure properties (e.g., polynomial-time, polylogspace, primitive recursive, etc.).

Remark 4.3. The reader might wonder why we will restrict ourselves to strict-type reductions, or slight variations, for the online setting. The reason is the following. Suppose that we have two (represented) online problems I_1 and I_2 . In an online way we want to use I_2 to solve I_1 . Now suppose that we have some online algorithm for I_2 . We could take a σ of length n representing an instance G_n of height n of I_1 , and convert it into an instance σ' of I_2 , and use it to produce a solution $s(\sigma')$ of I_2 , which could be converted back into a solution $s(\sigma) = A(s(\sigma'))$ of I_1 . The key issue we will investigate is how tight the relationships of sizes of the representations are. Ideally $|\sigma'| = |\sigma|$.

A problem $P = (I, O, s)$ is a *decision problem* if $O = \{0, 1\}$ and s is merely a predicate on I . This is the same as to say that any solution simply decides whether a predicate holds on a string or not. We say that $\sigma \in I$ is a *positive instance* of I if $s(\sigma) = 1$. Miltersen et.al. [52] analysed complexity classes for online algorithms, and in a slightly more general situation than our *monotone* one where, for example, the objects only get bigger. Miltersen et. al. [52] investigate online algorithms in which input data may change with time. For example, in a graph a vertex or an edge can disappear. Their reduction takes into account the potential changes of the input.

Definition 4.4. *Let C be a complexity class. A decision problem \mathcal{P} is C -incrementally reducible to another decision problem \mathcal{R} , denoted $\mathcal{P} \leq_{incr}^C \mathcal{R}$, if the following two conditions hold:*

1. *There is a transformation $T : I_{\mathcal{P}} \rightarrow I_{\mathcal{R}}$ in C which maps instances of \mathcal{P} to instances of \mathcal{R} such that $s_{\mathcal{P}}(\sigma) = s_{\mathcal{R}}(T(\sigma))$ (i.e., σ is a positive instance iff its image is a positive instance).*

2. There is a transformation Q in C which, given $\sigma \in I_{\mathcal{P}}$ and the incremental change δ to σ , where δ changes σ to σ' of the same length³, constructs the incremental change δ' to $T(\sigma)$ (where δ' changes $T(\sigma)$ to $T(\sigma')$).

Remark 4.5. We will here only consider C to be the class of polynomial time computable functions, and hence use \leq_{incr}^P accordingly. Milterson et. al. [52] also considered e.g C to be LOGSPACE. In [52] the authors specify the exact time bounds for all computations involved. This is the reason why they need the seemingly redundant part 2 of the definition above. Also, they look at auxiliary data structure generated for each instance and at the changes induced to the structure. However, from the perspective of general (e.g.) polynomial time computation this extra information is not necessary since these auxiliary bounds are evidently polynomial time.

The proposition below shows that $P \leq_{incr}^P Q$ is a variation of Weihrauch reduction from computable analysis which was independently rediscovered by computer scientists. Recall that strong Weihrauch reduction is witnessed by a pair of functionals Φ and Ψ .

Fact 4.6. *Suppose \mathcal{P} and \mathcal{Q} are online decision problems. Then $\mathcal{P} \leq_{incr}^P \mathcal{Q}$ iff $\mathcal{P} \leq_{sW}^P \mathcal{Q}$ with $\Psi = Id_{\{0,1\}}$.*

Proof. Suppose $P \leq_{incr}^P Q$. Then the transformation T from the definition of incremental reduction can be used as Φ in the definition of \leq_{sW}^P . Since σ is a positive instance iff $T(\sigma)$ is, $\Psi = Id_{\{0,1\}}$.

Conversely, suppose $P \leq_{sW}^P Q$ via $(\Psi, Id_{\{0,1\}})$, where Ψ is a polynomial functional from the space of inputs $I_{\mathcal{P}}$ of \mathcal{P} to the space of inputs $I_{\mathcal{Q}}$ of \mathcal{Q} . Then the first part of the definition of incremental reduction follows from the assumption that Ψ is a functional in C . By the continuity of Ψ and the fact that we used Id as the second functional, it suffices to deduce a polynomial time bound on the changes in the inputs of $\Psi(\sigma)$ based on the changes in σ . But this bound is just a big-O of the bound given by Ψ . \square

Following [52], we can impose specific bounds on the number of steps required for example, calculating δ' based on δ . The expectation is that it should be easier to make the change than to simply recompute $T(\sigma')$ “from scratch”. All these specialised bounds can also be expressed in terms of strong Weihrauch reduction; we omit details. As an application of Theorem 4.6 and various results in [52], we can obtain a number of polynomial time and polylogtime Weihrauch reductions in the study of online algorithms.

4.2. Weihrauch reduction and online graph colouring. Before we discuss the role of Weihrauch reduction in online colouring problem we give a brief overview of the latter.

4.2.1. Online graph colouring. Many problems can be re-cast as colouring problems, for example BIN PACKING. Indeed, colouring can be thought of as *avoiding configurations*. In basic graph colouring, we are simply avoiding an edge connecting vertices of the same colour, but we could instead avoid, for example, triangles or any finite set of configurations in some kind of constraint satisfaction problem.

³That is, δ is the difference between σ and σ' .

However, as this is an introductory paper we will stick to basic graph colouring. There is a large literature on this area such as Kierstead [44]. Graph colouring is quite a flexible tool, and many algorithmic meta-theorems such as for monadic second order logic (like Courcelle’s Theorem (see [16, 32])) can be viewed as colouring with constraints. We believe that this material has great online potential.

We will mention some of this in this subsection. As an illustrative example, we will online colour finite or infinite *trees* (or more accurately forests). So the objects of interest are trees being presented one vertex at a time so at height n we would have a forest of n vertices. Then we can use a compact representation $T \subset \omega^\omega$ of the collection of forests of height n . The result below is an easy (restated in our notation) result from the folklore essentially following from Bean [9]. It works for any (not necessarily primitive recursive) general computable online procedure.

Proposition 4.7. *For every online algorithm A there is a σ in T of length 2^{t-1} such that the tree A acting on $T_{F(\sigma)}$ needs at least t colours.*

We will write $\chi_A(G_\sigma)$ for the number of colours used to colour G when processed by online algorithm A . The above is nearly optimal, in that we have the following:

Theorem 4.8 (Lovasz, Saks and Trotter [50]). *There exists an online algorithm A such that for every 2-colourable graph G , if G has n vertices then $\chi_A(G) \leq 1 + 2 \log n$.*

This brings us to the notion of a *performance ratio*. Most combinatorial algorithms we would teach in a standard combinatorics class are *offline*. This means that for a finite structure H , say, the algorithm has H as part of the input and calculates using the global structure of H .

Consider the situation of an inductive problem in a class \mathcal{C} , and suppose we have an optimisation problem. Then associates with a $\sigma \in I, \tau \in S$ will be a *cost* function $c(\sigma, \tau)$. This is the cost of solution τ for problem instance σ . Now the *performance ratio* of an algorithm f is the ratio of $c(\sigma, o(\sigma))$ with $c(\sigma, f(\sigma))$ where $o(\sigma)$ is an optimal solution for σ ; the offline solution.

We illustrate with colourings. The *offline* chromatic number of G_σ will be denoted by $\chi_{\text{off}}(G_\sigma)$ and for forests, we would have

$$\chi_{\text{off}}(T_\sigma) = 2.$$

Definition 4.9 (Sleator and Tarjan [58]). The *performance ratio* is defined as be

$$r(\sigma) = \frac{\chi_A(T_\sigma)}{\chi_{\text{off}}(T_\sigma)}.$$

Here we are stating the definition for colouring but the definition applies to any online optimisation problem, as above. In the case of colouring forests, we see that the approximation ratio is $O(\log(|\sigma|))$. In the infinite case, the relevant approximation ratio is the growth rate of $r(\sigma)$ for all paths in the tree T representing the problem.

For example, a graph is called *d-inductive* (or *d-degenerate*) if the vertices of G can be ordered as $\{v_1, \dots, v_n\}$ so that for every $i \leq n$, $|\{j > i \mid v_i v_j \in E\}| \leq d$. For example, by Euler’s formula, all planar graphs are 5-inductive. For those who know some graph theory, *d-inductive* graphs also include all graphs of treewidth d , and extremely important class in algorithmic graph theory (see Downey and Fellows [16], for example). Again note that *d-inductive* graphs have a compact representation (space).

Theorem 4.10 (Irani [33, 34]). *Let σ represent a d -inductive graph of height n . Then first fit will use at most $O(d \log n)$ many colours to colour G_σ . Moreover, for any online algorithm A , there is a d -inductive G_σ such that $\chi_A(G_\sigma)$ is $\Omega(d \log n)$.*

Sometimes, this growth rate reaches a limit, as in problems with *constant* approximation ratios.

The classical example is BIN PACKING. We can think of bins as colours, and the objects having sizes and the constraint being that we cannot have more objects of a specific colour than the bin constraint. That is, BIN PACKING takes as input sizes $a_i \in \mathbb{N}$ and a parameter V for simplicity, and colours a_i with colour $c(a_i)$ subject to $\sum_{c(a_i)=c} a_i \leq V$, and seeks to minimize the number of colours.

Theorem 4.11 (see [27]). *First fit gives a performance ratio of 2 for online BIN PACKING.*

Notice that BIN PACKING is another example of colouring with constraints.

4.2.2. Online reduction. In this subsection we define a new version of Weihrauch reduction, and we also give a non-trivial example of such a reduction between two distinct online problems.

For convenience we use 2^ω as the ambient totally disconnected space, but otherwise use appropriate names. Let X and Y be spaces represented by 2^ω . Again we think of f and g being solutions for minimisation X and Y -problems respectively. Thus, for example, we are thinking of X and Y as inductive structures with filtrations $\{X_n \mid n \in \mathbb{N}\}$ and $\{Y_n \mid n \in \mathbb{N}\}$ respectively. Then the strings of length n represent the structures of height n , and $f(\sigma)$ will represent a solution to the problem represented by σ . Thus they will have an associated *cost* which in the case of colouring is the number of colours, denoted $c(\cdot)$. We will denote f_{off} and g_{off} as offline solutions. That is $f_{\text{off}}(\sigma)$ would be the solution to the minimisation problem X_n of height n with $\delta(\sigma) = X_n$, and similarly g_{off} .

We state the below for single valued functions, but again there is an analogous multi-valued version, where the solution produced for g should be within the correct ratio. The idea of the following is that on input $\alpha \upharpoonright n$, we want to compute (a representation of) $f(\alpha \upharpoonright n)$ ⁴ To to this we will apply (a representation of) B to generate an input to (a representation of) an input for g , and then use the algorithm A to translate this back to give $f(\alpha \upharpoonright n)$. Again we emphasize that this is all working with representations, and should be read this way.

Definition 4.12. Let f, g be functions on 2^ω . Then f is called *ratio preserving online reducible to g* , $f \leq_O^r g$, if there are (type II) online computable functions A and B with and a constant d , such that for all n ,

$$f(\alpha \upharpoonright n) = A(\alpha \upharpoonright n, g(B(\alpha \upharpoonright n))),$$

and the ratio of $c(f(\alpha \upharpoonright n))$ to $c(f_{\text{off}}(\alpha \upharpoonright n))$ is at most d times the ration of $c(g(B(\alpha \upharpoonright n)))$ to $c(g_{\text{off}}(B(\alpha \upharpoonright n)))$.

The fact below isolates the most important feature of the reduction.

Fact 4.13. *If $f \leq_O^r g$ then , for some $d > 0$,*
$$\frac{c(f \upharpoonright n)}{c(f_{\text{off}} \upharpoonright n)} \leq d \frac{c(g \upharpoonright n)}{c(g_{\text{off}} \upharpoonright n)}.$$

⁴We want to avoid explicit representations, but of course we should have F representing f with F acting on 2^ω , and for any $\alpha \in 2^\omega$, $\lim_n F(\alpha \upharpoonright n)$ realizes (represents) $f(\alpha)$.

To give a non-trivial example of an online reduction we need several definitions.

In classical colouring, Kierstead investigated online colouring of INTERVAL GRAPHS. A graph $G = (V, E)$ is called a k -interval graph if each vertex v of G can be represented by a closed subinterval of $[0, 1]$ such that if I_v represents v and I_w represents w , then if $vw \in E$, $I_v \cap I_w \neq \emptyset$, such that the largest number of intersecting intervals (the *cutwidth*) is $\leq k$. These are exactly the graphs which have $Pathwidth \leq k$, a graph metric coming from the Robertson-Seymour minors project (see [56, 16]).

Definition 4.14. Let $ColInt_k$ denote the online problem of colouring a k -interval graph. (We leave the precise representation of the problem to the reader.)

The other online problem is on covering of an interval partial ordering by chains. A partial ordering (P, \leq) is called an *interval ordering* if P is isomorphic to (I, \leq) where I is a set of intervals of the real line and $x \leq y$ iff the right point of x is left of the left point of y . Interval orderings can be characterised by the following theorem.

Theorem 4.15 (Fishburn [24]). *Let P be a poset. Then the following are equivalent.*

- (a) P is an interval ordering.
- (b) P has no subordering isomorphic to $\mathbf{2} + \mathbf{2}$ which is the ordering of four elements with $\{a, b, c, d\}$ with $a < b$, $c < d$ and no other relationships holding.

The width of an interval ordering (P, \leq) is defined naturally to be the minimum over all presentations of the maximum number of intervals covering some point of $[0, 1]$. Given an interval ordering (P, \leq) of width k , our goal is to cover it with as few chains as possible; the chains do not have to be disjoint.

Definition 4.16. Let $ChInt_k$ denote the online problem of covering an interval ordering (P, \leq) of width k by not necessarily disjoint chains. (We leave the precise representation of the problem to the reader.)

The theorem below gives a non-trivial example of an online ratio-preserving reduction between online problems. The proof of the theorem below is essentially an analysis of the clever argument given in Kierstead and Trotter [45].

Theorem 4.17. *For any positive $k \in \mathbb{N}$ there is an online solution g to $ChInt_k$ with a constant performance ratio which can be transformed into an online solution f to $ColInt_k$ with the property $f \leq_O^r g$ via a constant $d = 1$.*

Corollary 4.18 (Kierstead and Trotter [45]). *There is an online algorithm to colour k interval graphs with a constant competitive ratio.*

Proof of Corollary. Kierstead and Trotter [45] showed that every (P, \leq) (online) interval ordering of width k can be online covered by $3k - 2$ many chains. Here recall that a *chain* in a partial ordering is a \leq -linearly ordered subset. A collection of chains $\{C_1, \dots, C_q\}$ covers (P, \leq) if each element of P lies in one of the chains. An *antichain* is a collection of pairwise \leq -incomparable elements. We will see that $ColInt_k \leq_O^r ChInt_k$ is witnessed via a reduction with constant $d = 1$. It remains to apply Fact 4.13. \square

Proof of Theorem 4.17. The basic idea is quite simple. Take our online k interval graph, turn it into an online interval ordering of width k , and then consider that

chain covering as a colouring. However, to see that this idea works, we need to argue that there is an online solution g to the interval chain covering problem which uses only the information about comparability of various elements, and not their ordering.

We first prove the following. Suppose that (P, \leq) is a online interval ordering of width k . Then P can be online covered by $3k - 2$ many chains. We need the following lemma whose proof is fairly straightforward. For a poset P , and subsets S, T , we can define $S \leq T$ iff for each $x \in S$ there is some $y \in T$ with $x \leq y$. (Similarly $S|T$ etc.)

Lemma 4.19. *If P is an interval order and $S, T \subset P$ are maximal antichains then either $S \leq T$ or $T \leq S$.*

The algorithm for chain covering uses induction on k . We consider the vertices as $1, 2, \dots$ with p added at step p . If $k = 1$ then P is a chain, and there is nothing to prove. Suppose the result for k , and consider $k = 1$. We define B inductively by

$$B = \{p \in P : \text{width}(B^p \cup \{p\}) \leq k\}.$$

Here B^p denotes the amount of B constructed by step p of the online algorithm. Then B is a maximal subordering of P of width k . By the inductive hypothesis the algorithm will have covered B by $3k - 2$ chains. Let $A = P - B$. Now it will suffice to show that A can be covered by 3 chains, and then these will be covered by the greedy algorithm.

To see this it is enough to show that every elements of A is incomparable with at most two other elements of A . Then the greedy algorithm will cover A , as we see elements *not* in B .

Lemma 4.20. *The width of A is at most 2.*

Proof. To see this, consider 3 elements $q, r, s \in A$. Then there are antichains Q, R, S in P of width k with $q|Q, r|R$ and $s|S$. Moreover these can be taken as maximal antichains. Applying Lemma 4.19, we might as well suppose $Q \leq R \leq S$. Suppose that $r|q$ and $r|s$. Then we prove that $q < s$. Since $q|r$ and $\text{width}(P) \leq k + 1$, there is some $r' \in R$ with q and r' comparable. Since $q|Q, r' \notin Q$. Since the width of B is $\leq k$, there is some $q' \in Q$ q' and r' comparable. Since $Q \leq R$, there is some $r_0 \in R$ with $q' \leq r_0$. Since eR is an antichain, $q' \leq r'$. Since $q|q', q \leq r'$. Similarly, there exists $r'' \in R$ with $r'' \leq s$. Since P does not have any ordering isomorphic to $\mathbf{2} + \mathbf{2}$, we can choose $r' = r''$, and hence $q < s$. \square

Now we suppose that r, q, s, t are distinct elements of A with $q|\{r, s, t\}$. Then without loss of generality $r < s < t$ since the width of A is at most 2. Since $s \in A$ there is an antichain $S \subset B$ of length k with $s|S$. Since $s|q$, and $\text{width}(P) \leq k + 1$, q is comparable with some element $s' \in S$. If $s' < q$, then $s'|r$ and hence the suborder $\{s', q, r, s\}$ is isomorphic to $\mathbf{2} + \mathbf{2}$. Similarly, $q < s'$ implies $s'|t$ and then the subordering $\{q, s', s, t\}$ is isomorphic to $\mathbf{2} + \mathbf{2}$. Thus there cannot be 4 elements r, q, s, t of A with $q|\{r, s, t\}$. Hence A can be covered by 3 chains.

It is easily see that the procedure above uses only comparability of intervals. Thus, the theorem follows. \square

Problem 4.21. *Investigate online reduction between online algorithms in the literature.*

We also expect that the online reduction may lead to new online algorithms based on the already existing ones.

Also graphs with constrained decompositions such as those of bounded treewidth, pathwidth, clique-width, etc have been extensively studied in the literature, and particularly combine well with algorithmic meta-theorems (see e.g. Downey-Fellows [16], Flum and Grohe [25], Grohe [32] for a sample).

One example is given by k -interval graphs met above which are those of pathwidth $\leq k$. A G of pathwidth k has a *path decomposition* which is a collection of sets of vertices V_1, \dots, V_n all of size $\leq k + 1$ such that for all vertices $v \in V(G)$, there is at least one i with $v \in V_i$, if $xy \in E(G)$, then for some i , $\{x, y\} \subseteq V_i$ and finally if $x \in V_i$ and $x \in V_j$ (with $i < j$) then for all $q \in [i, j]$, $x \in V_q$. The last property is called the *interpolation property*, and says that pathwidth is kind of a measure of how far you are from being either a grid or a clique.

Now *given* such a path decomposition, and some optimisation property we want to solve (such as for the largest clique), if the property is definable in monadic second order logic (even with counting), then we can solve the problem by dynamic programming (actually using special automata) beginning at V_1 and finishing at V_n by the methods of Courcelle [16, 25, 32].

Problem 4.22. *Investigate the extent to which this dynamic programming is online. Presumably, it will be online for properties defined by monadic second order counting logic with counting modulo some kind of delay.*

Moreover, as we have seen above for the special case of colouring above, we get a constant ratio approximation algorithm, for a graph of pathwidth k , no matter how we are given the online presentation. The difference is that if we are given a path decomposition as the presentation, then $k + 1$ colours will suffice. But perhaps the methods for colouring are more general. The point is that graphs of bounded pathwidth have very constrained structure.

Problem 4.23. *Investigate the approximability of monadic second order definable properties on graphs of bounded pathwidth, but given as arbitrary online presentations.*

The same can be asked for graphs of bounded treewidth which has the same definition as pathwidth, but the structure of the decomposition is a tree and not a path. These also have dynamic programming algorithms, but are always *leaf to root*, whereas even given a tree decomposition as an online root to leaf structure, presumably some kind of algorithm will work, but it will no longer be automatic. This seems a great area to pursue.

Also related seems the idea of online parameterised problems [20, 21], where we want an online solution to a problem with a fixed parameter. For example, k -VERTEX COVER asks for a collection of vertices where each edge of a graph includes at least one of the vertices, and this is polynomial time for a fixed k . This is also online polynomial time for a fixed k by the following simple algorithm (so long as we are allowed 2^k many possible solutions). We can use the following simple method of building a tree of height 2^k by taking an edge, and branching on that edge, and then deleting the covered edges, and repeating. This process is also online. In [20, 21] Downey and McCartin showed that the online view brings to the other parameters such as what they call *persistence* which characterises the extent to which a path decomposition does not resemble a fuzzy ball. The point is that online algorithms

point at new parameters of a problem which deserve attention, in the same way that parameterised complexity showed that parameters allow a more fine grained understanding of the computational complexity of a combinatorial problem.

5. Δ_2^0 PROCESSES, FINITE REVERSE MATHEMATICS, AND WEIHRAUCH REDUCTION

Imagine we are in a situation where the data we are dealing with is so large that we cannot see it all. At each stage s our goal is to build a solution f to some problem. But there might be no hope of giving a fixed solution at each stage n , and like a Triage Nurse making an ordering for patients to obtain medical attention, we would update our solution as more information becomes available. So for each $n \leq s$ we would be computing $f(n, s)$ from the finite information σ with $|\sigma| = n$. For simplicity we state the next definition for combinatorial problems with totally disconnected representations, and take 2^ω as the representing example.

Definition 5.1. A limiting online algorithm on 2^ω is a computable function A such that for each s , $A(\alpha \upharpoonright s)$ computes a string $\{f_A(n, s) \mid n \leq s\}$ such that $\lim_s f_A(n, s)$ exists for each n .

As usual we would have $A(\alpha \upharpoonright g(s))$ for the g -online version.

We can then compare combinatorial problems by how fast their limits converge.

Definition 5.2. We say that algorithm $A \leq_{O,lim} B$ if there is an online Weihrauch reduction of A to B such that $f_B(n, s) = f_B(n, t)$ for all $t \geq s$ implies $f_A(n, s) = f_A(n, t)$ for all $t \geq s$.

This gives a fine grained measure of the complexity of combinatorial problems. For example, consider the “theorem” that every finite binary tree of height n has a path of length n . Then we can consider the existence of a uniform function A which takes a given binary tree of height n to a path. This is an online limit problem where the underlying space X is that with nodes generated by the collection of binary trees of height n at level n . The completion of this will represent paths through infinite binary trees.

Remark 5.3. We could argue that the Reverse Mathematics principle WKL_0 which states that every infinite binary tree has a path, is equivalent to the statement that there is a limiting online algorithm for finding paths which works on X . We call this *limiting online paths*.

A binary tree T of height n is called *separating* if for each $j \leq n - 1$, for any node σ on T of height j , and $i \in \{0, 1\}$, if $\sigma * i$ does not have an extension in T of height n , then for all τ of length j , neither does $\tau * i$. Let X_S be the totally disconnected space representing the collection of all separating finite trees. The following is an online interpretation and refinement of the classical fact that Weak König’s Lemma is equivalent to Weak König’s Lemma for separating classes.

Proposition 5.4. There is a $(2^{n+1} - 2)$ -limiting online reduction which finds limiting online paths in X from those in X_S .

Proof. We remind the reader of how this proof works. Suppose we have a tree T_s of height s . In an online fashion, we will generate a tree H of height 2^{s+1} . This is done inductively. At step 1, we can think of the nodes labeled 0 and 1 in T as being represented by 0 and 1 in H . At step 2, in T it is possible for us to have

00, 01, 10, 11 and these are represented by 4 levels in H , with height 2 representing 00, level 3 01, level 4 10, and level 5 11. Now we continue inductively. This makes level n of T correspond to trees of height $2 + 4 + \dots + 2^n = 2^{n+1} - 2$. As the construction proceeds, if some σ fails to have an extension at length s , in T_s , there will be some shortest $\sigma' \preceq \sigma$ which fails to have a length s extension in T_s . Then in H_s , we don't extend to length s (from length $s - 1$) all paths corresponding to $\nu * j$ with j representing σ' in H_{s-1} .

Consider any limiting online algorithm for finding a path for path α corresponding to H , in X_0 . This naturally and in an online way allows us from level 2^{s+1} to generate an online path in T_s , and is clearly a limiting online reduction. \square

Problem 5.5. *Figure out the smallest g in place of 2^{s+1} in the reduction above, which would give a precise measure of how tight the reverse mathematics relationship is.*

There seems a whole research programme available here. For example, we could be given an online bipartite graph B_σ for $\sigma \prec \alpha$. We either have to build a complete matching or demonstrate that Hall's condition fails. One representation of this problem will involve a compact space where the nodes are bipartite graphs of height $2n$, say, and where the paths all represent graphs which obey Hall's condition. The online operator will act on this compact tree of representations for graphs B_σ . Now as the process goes along, we might have to update the solution at hand. That is, the online process has $B_\sigma \mapsto M_\sigma$,

One intriguing example is that of finding a basis in a vector space. In the case that the vector space is over the rationals, then presumably this will correlate to some principle like ACA_0 . But consider a finite field such as $\text{GF}(2)$. We know that RCA_0 proves that we can find a basis for a vector space for this field. But it is not hard to construct an online vector space over $\text{GF}(2)$ for which there is no online algorithm to do this, unless we have a computable delay. Comparing the online complexity of such problems with such computable delay would see to give significant insight into the fine structure of reverse mathematics. In this particular case, we also note that a polynomial time algorithm for finding a basis of a polynomial time vector space was proven to be equivalent to $P = NP$ suggesting intriguing connections with complexity theory.

We remark that there are many processes that have been investigated and fall under the model we have introduced. One such example is algorithmic learning theory, such as EX -learning (Gold [31]). Here one is presented with a_0, a_1, \dots values for a function $f(0), f(1), \dots$, and we need to eventually print out an index for $\varphi_e = f$ from some point onwards. This is clearly an example of an online algorithm, and fits into this section as a limiting algorithm. Perhaps there are connections with this and reverse mathematics, and this remains to be explored. Another area which could be incorporated would be asynchronous computing. Here we have a series of agents A_1, \dots, A_k communicating through asynchronous channels, and attempting to compute a set of functions f_1, \dots, f_k , where there might be e.g. some kind of crash failure meaning that one of the agents dies and stops sending signals. For example, the CONSENSUS problem asks for all the f_i 's which have not crashed to give the same value. A *run* could be represented in a space of possible communications and failures. There are a number of reductions which have been produced in this area, showing that Consensus is a certain kind of minimal failure, and other problems can be solved if Consensus can (Chandra and Toueg [14]). It

would be interesting to see if these results can be placed in the hierarchy of online limiting reductions, since they appear to look like online limiting reductions.

Finally, one exciting possibility would be to include randomization in this setting. Randomized online algorithms are quite common in practice (see e.g. Albers [4]). For us we could use the theory of algorithmic randomness (see [17, 49, 54]) easily. For example, an online algorithm with randomized advice (i.e. representing a coin toss at each stage) could be done via (using 2^ω as a representative space) by considering online algorithms from $2^\omega \times 2^\omega \rightarrow S$, with S some solution space, with the first copy of 2^ω representing the problem, the second representing “advice” strings and S the solution space. The online algorithm could take $(\sigma, \tau) \rightarrow s_n$, and would run on extensions of τ provided that $[\tau]$ avoids some algorithmic randomness test, such as a Martin-Löf test. Using oracles we could also tie this to the theory of algorithmic randomness using the “*fireworks*” method of Shen (see Bienvenu and Patey [10]). These ideas remain to be explored.

6. REAL FUNCTIONS.

So far all objects of study have been discrete and spaces compact. However, there is a perfectly reasonable extension of these ideas to continuous objects such as the space of continuous functions on the unit interval. There has been a lot of work on complexity theory of real functions; see, e.g., Ko [47]. In terms of applications, a natural object of study would be online analysis; analytic processes which run quickly and only use local knowledge of the precision of the inputs. As we observe below with natural representations, addition of reals x, y with precision 2^{-n} only needs x and y to within $2^{-(n+2)}$. Integration and other standard processes have similar commentry, but we leave this to a later paper. Also there are other online processes on non-compact spaces, such as EX-learning, or the KC theorem discussed earlier. We also defer discussion of such topics for later papers, and here stick to analysis.

The main goal of this section is to demonstrate the role of primitive recursion as a useful abstraction. The content of this section is not technically hard, but one can easily imagine a much deeper general framework that could emerge from these basic ideas.

Recall that a Cauchy sequence $(r_i)_{i \in \mathbb{N}}$ of rationals is fast if $|r_i - r_{i+1}| < 2^{-i}$, for every i . These are the *names* which represent the space. A function $f : [0, 1] \rightarrow \mathbb{R}$ is *computable* if there is a Turing functional Φ such that, for each $x \in [0, 1]$ and for every fast Cauchy sequence χ converging to x , the functional Φ enumerates a fast Cauchy sequence for $f(x)$ using χ as an oracle. In particular, using the terminology, we would be generating a representation of the function via names of Cauchy sequences in such a way that it is representation independent. That is, $(\Phi^\chi(n))_{n \in \mathbb{N}}$ is a fast Cauchy sequence for $f(x)$. This in particular means that, on input $(r_i)_{i \in \mathbb{N}}$, the use of $\Phi^{(r_i)_{i \in \mathbb{N}}}(j)$ corresponds to δ when $\epsilon = 2^{-j+1}$ in the standard ϵ - δ definition of a continuous function.

It is well-known that Weierstrass approximation theorem is effectivisable in the sense of Turing computability [55]. This means that a function $f : [0, 1] \rightarrow \mathbb{R}$ is computable iff there is a computable sequence of polynomials $(p_i)_{i \in \mathbb{N}}$ with rational coefficients with the property

$$\sup_{x \in [0, 1]} |f(x) - p_i(x)| < 2^{-i},$$

for every i .

We have seen that the most general definition of being online for combinatorial structures involves being g -online for some primitive recursive function g . That is, there is a translation between using $g(n)$ many bits of α to compute n bits of $f(\alpha)$. We have also seen that for most natural online situations, we can translate this to a wider tree where $\alpha' \upharpoonright n$ represents $\alpha \upharpoonright g(n)$, so we can use strict (ibT primitive recursive) procedures. It is not completely clear if this is natural in the setting of analysis, since we might wish to stick to *standard* representations of the spaces, like 2^ω and ω^ω , as above.

We first consider the most general setting where we allow g -online for a primitive recursive g , so using $g(n)$ bits to decide the output for length n . We will call this *punctually* computable. In this case, there are two natural definitions of what it would mean for such an f to be “online” computable in the most general sense of primitive recursion. The first notion is the most straightforward sub-recursive version of the standard definition.

Definition 6.1. A function $f : [0, 1] \rightarrow \mathbb{R}$ is *punctually computable* if there is a primitive recursive functional Φ such that, for each $x \in [0, 1]$ and for every fast Cauchy sequence χ converging to x , the functional Φ enumerates a fast Cauchy sequence for $f(x)$ using χ as an oracle.

By restricting ourselves to dyadic rationals, we can assume that fast Cauchy sequences come from a compact totally disconnected space of the names of dyadic rationals in $[0, 1]$. Thus, Lemma 2.5 can be applied to ensure that there is no ambiguity in the notion of a primitive recursive functional in this case. In particular, the definition has a natural polynomial-time version which we omit (see [47]); the same applies to any natural complexity class which may be of interest.

The second version filters through the theorem of Weierstrass. It views f as a primitive recursive point in the metric space $(C[0, 1], \text{sup})$ rather than as a functional.

Definition 6.2. A function $f : [0, 1] \rightarrow \mathbb{R}$ is *uniformly punctually computable* if there is a primitive recursive function which on input i outputs (the index of) a polynomial p_i with rational coefficients such that $\sup_{x \in [0, 1]} |f(x) - p_i(x)| < 2^{-i}$.

Clearly, there is a natural polynomial-time modification of the definition above which we omit.

Every uniformly punctually computable f is punctually computable. Are these two definitions equivalent? It is not completely evident why Weierstrass approximation theorem should hold primitively recursively. Indeed, in the standard Turing computable proof we would wait for a cover of $[0, 1]$ by δ_i -balls B_i such that $f(B_i)$ has diameter $< \epsilon$, for every i . It seems that even when f is punctual this search could be unbounded.

Nonetheless, the theorem below shows that these definitions are equivalent. This result is not really new. With some effort its proof can be extracted from book [47], but the book is mainly focused on polynomial time and exponential versions of the definitions above. There is much combinatorics specific to complexity theory which significantly obscures the idea behind the proof. Primitive recursion strips away complex counting combinatorics thus clarifying the idea.

Theorem 6.3. *Every punctually computable $f : [0, 1] \rightarrow \mathbb{R}$ is uniformly punctually computable.*

Proof sketch. The idea here is similar to that in the proof of Lemma 2.5. Fix n and consider the functional $\Psi_n^x = \Phi^x(n)$ which uniformly primitively recursively outputs the first few bits of $f(x)$ up to error 2^{-n} , for any input x . Since Ψ_n is given a primitive recursive scheme (with parameter n), we can work by induction on the complexity of the scheme and emulate all its possible computations at once, as in Lemma 2.5. Since the space of dyadic presentations of rationals is primitively recursively compact, this will lead to a primitively recursively branching tree of possible computations whose height is determined by the syntactical complexity of the primitive recursive scheme. By the choice of Ψ_n , one of these computations must work for an arbitrary $x \in [0, 1]$. Thus, we have primitively recursively calculated an open cover $[0, 1]$ by basic open intervals J_1, \dots, J_k , such that whenever $x, y \in J_i$ we have $|f(x) - f(y)| < 2^{-n+1}$. If z_i is the center of J_i , then define (the graph of a) piecewise linear function h_n by connecting points $(z_i, \Psi_n^{z_i})$ and $(z_{i+1}, \Phi_n^{z_{i+1}})$, $i = 1, \dots, n-1$. Note that the values of the $\Phi_n^{z_i}$ have already been calculated. Since the intervals are overlapping, this piecewise linear function h_n approximates f with precision 2^{-n+2} . We can primitively recursively smoothen h_n by replacing it with a polynomial p_n such that $\sup_{x \in [0,1]} |p_n(x) - f(x)| < 2^{-n+3}$. \square

See Chapter 8 of [47] for a detailed analysis of the polynomial-time versions of Weierstrass approximation theorem. Recall that in the proof sketch above we generated the tree of possible computations. For a polynomial-time operator this tree may be exponentially large at worst. This difficulty cannot be circumvented and the polynomial-time analogy of the theorem above *fails* as explained in great detail in [47].

We see that punctual analysis fits somewhere in-between computable analysis and polynomial-time analysis, and there is likely much depth in the subject. Such a theory could provide us with a stronger technical link between computable and feasible analysis. Nonetheless, it seems there has been no dedicated study of primitive recursive continuous functions and punctual presentations of analytic spaces.

Problem 6.4. *Develop primitive recursive (“punctual”) analysis.*

Now in the case that we want to look at the strictly online model, we are stuck with using, for instance, the bit representation of a real x , and would be working, for example, with 2^ω . Then to compute $f(x)$ with precision 2^{-n} we would need $x \upharpoonright n$. We might ask for delay k so might use $2^{-(n+k)}$. Now in this case, we see that, for example addition is online (on $2^\omega \times 2^\omega$) with delay 2, and if f is a given online computable function which is bounded then $\int_0^x f(x)dx$ would also be online computable with delay 2. We remark that this model would seem to be one emulating classical numerical analysis. This fine-grained analysis of what is actually online seems completely open and is clearly related to numerical analysis.

REFERENCES

[1] P. E. Alaev. Structures computable in polynomial time. I. *Algebra Logic*, 55(6):421–435, 2017.
 [2] P. E. Alaev. Structures computable in polynomial time. II. *Algebra Logic*, 56(6):429–442, 2018.
 [3] Alaev, P., and Selivanov, V., Polynomial-Time Presentations of Algebraic Number Fields. *Computability in Europe 2018*, (2018) 20-29.

- [4] Albers, S., Online Algorithms: A Survey, in *Mathematical Programming*, Vol. 97, (2003) 3–26.
- [5] Barmpalias, G. and Lewis A., A c.e. real that cannot be sw-computed by any Ω -number, *Notre Dame Journal of Formal Logic*, 47 (2006), 197–209.
- [6] Bazhenov, N., Downey, R., Melnikov A., and Kalimullin, I. Foundations of online structure theory, *Bulletin of Symbolic Logic*, Vol. 25 (2019), 141–181.
- [7] Bazhenov, N., Harrison-Trainor, M., Kalimullin, I., Melnikov A., and Ng, K.M. Automatic and polynomial-time algebraic structures, *J. Symbolic Logic*. In press.
- [8] Allan Borodin and Ran El-Yaniv. *Online computation and competitive analysis*. Cambridge University Press, New York, 1998.
- [9] Bean, D., Effective Coloration, *J. Symbolic Logic* Volume 41, Issue 2 (1976), 469–480.
- [10] Bienvenu, L., and Patey, L., Diagonally non-computable functions and fireworks, submitted. (arXiv:1411.6846)
- [11] Brattka, V. and Gherardi, G., Weihrauch Degrees, Omniscience Principles and Weak Computability, in: Bauer, Andrej and Hertling, Peter and Ko, Ker-I (eds.), *CCA 2009, Proceedings of the Sixth International Conference on Computability and Complexity in Analysis*, Leibniz-Zentrum für Informatik, Schloss Dagstuhl, Germany, 2009, pages 83–94.
- [12] Braverman, M., On the Complexity of Real Functions *Foundations of Computer Science, 2005, FOCS'05*.
- [13] D. Cenzer and J. B. Remmel. Complexity theoretic model theory and algebra. In Yu. L. Ershov, S. S. Goncharov, A. Nerode, and J. B. Remmel, editors, *Handbook of recursive mathematics, Vol. 1*, volume 138 of *Stud. Logic Found. Math.*, pages 381–513. North-Holland, Amsterdam, 1998.
- [14] Chandra T. D. and Toueg, S. Unreliable failure detectors for reliable distributed systems. *JACM* 43(2):225–267, 1996.
- [15] Csima, B., *Applications of Computability Theory to Prime Models and Differential Geometry*, Ph.D. dissertation, The University of Chicago, 2003.
- [16] Downey, R. and Fellows, M. *Fundamentals of Parameterized Complexity*, Springer-Verlag, 2013.
- [17] Downey, R., and Hirschfeldt, D., *Algorithmic Randomness and Complexity*, Springer-Verlag, 2010.
- [18] Downey, R., Harrison-Trainor, M., Kalimullin, I., Melnikov, A., and Turetsky, D., Graphs are not universal for online computability, submitted.
- [19] Downey, R., LaForte, G., and Hirschfeldt, D., Randomness and reducibility, extended abstract appeared in *Mathematical Foundations of Computer Science, 2001* J. Sgall, A. Pultr, and P. Kolman (eds.), *Mathematical Foundations of Computer Science 2001, Lecture Notes in Computer Science 2136* (Springer, 2001), 316–327). final version in *Journal of Computing and System Sciences*. Vol. 68 (2004), 96–114.
- [20] R. Downey and C. McCartin, Online problems, pathwidth, and persistence, *Proceedings IWPEC 2004*. Springer-Verlag Lecture Notes in Computer Science 3162, pp 13–24, 2004.
- [21] R. Downey and C. McCartin, Online promise problems with online width metrics, *Journal of Computing and System Sciences*, Vol. 73, No. 1 (2007), 57–72.
- [22] David B. A. Epstein, James W. Cannon, Derek F. Holt, Silvio V. F. Levy, Michael S. Paterson, and William P. Thurston. *Word processing in groups*. Jones and Bartlett Publishers, Boston, MA, 1992.
- [23] Fiat, A., and Woeginger, G., *Online Algorithms*, Springer LNCS, vol 1442.
- [24] Fishburn, P., Interval graphs and interval orders, *Discrete Mathematics*, Volume 55, Issue 2, July 1985, Pages 135–149.
- [25] Flum, J., and Grohe, M., *Parameterized Complexity Theory*, Springer-Verlag, 2006.
- [26] Friedman, H. and Ko, K, Computational complexity of real functions. *Theoret. Comput. Sci.*, 20(3):323–352, 1982.
- [27] Garey, R. and Johnson, D., *Computers and Intractability*, W. H. Freeman, 1979.
- [28] Gasarch, W., A survey of recursive combinatorics, in *Handbook of Recursive Mathematics*, Vol. 2., Edited by Ershov, Goncharov, Marek, Nerode, and Remmel. 1998. Pages 1041–1176. Published by Elsevier ISBN: 0444544249.
- [29] Grzegorzcyk, A., Computable functionals, *Fundamenta mathematica*, Vol. 42 (1955), 168–202.
- [30] Serge Grigorieff. Every recursive linear ordering has a copy in $DTIME-SPACE(n, \log(n))$. *J. Symb. Log.*, 55(1):260–276, 1990.

- [31] Gold, M., Language Identification in the Limit, *Information and Control*, Vol. 10 (1967), 447–474.
- [32] Grohe, M., *Descriptive Complexity, Canonisation, and Definable Graph Structure Theory*, Lecture Notes in Logic, Volume 47. Cambridge University Press, 2017
- [33] Irani, S., Coloring Inductive Graphs On-Line, in *Proceedings for for the 31st Symposium on the Foundations of Computer Science*, 1990, pp. 470–479.
- [34] Irani, S., Coloring Inductive Graphs On-Line, *Algorithmica*, vol.11, no.1, Jan. 1994, pp.53-72.
- [35] Kalimullin, I., Melnikov, A., and Montalban, A., Punctuality on a cone. In preparation.
- [36] Kalimullin, I., Melnikov, A., and Ng, KM., The Diversity of Categoricity Without Delay. *Algebra and Logic*. 56(2) (2017), 171-177.
- [37] Kalimullin, I., Melnikov, A., and Ng, KM., Algebraic structures computable without delay. *Theoretical Computer Science*. 674 (2017), 73-98.
- [38] Karp, R., Reducibility Among Combinatorial Problems In R. E. Miller; J. W. Thatcher (eds.). *Complexity of Computer Computations*. New York: Plenum. (1972), 85-103.
- [39] Karp, R., On-line algorithms versus off-line algorithms: How much is it worth to know the future? IFIP Congress (1). 12 (1992), 416-429.
- [40] Kawamura, A., and Cook S., Complexity Theory for Operators in Analysis, *ACM Transactions on Computation Theory*, 4(2), Article 5, 2012.
- [41] Khoussainov, B., A quest for algorithmically random infinite structures. *Proceedings of LICS-CSL 2014 conference*. Vienna, Austria.
- [42] Bakhadyr Khoussainov and Anil Nerode. Automatic presentations of structures. In *Logic and Computational Complexity (Indianapolis, IN, 1994)*, volume 960 of *Lecture Notes in Comput. Sci.*, pages 367–392. Springer, Berlin, 1995.
- [43] Kierstead, H., An Effective Version of Dilworth’s Theorem, *Trans. Amer. Math. Soc.*, 268(1)-November 1981.
- [44] Kierstead, H., Recursive and On-Line Graph Coloring, In *Handbook of Recursive Mathematics*, Volume 2, pp 1233-1269, Elsevier, 1998.
- [45] Kierstead, H. and Trotter, W., An Extremal Problem in Recursive Combinatorics. *Congressus Numeratum*, 33, pp 143-153, 1981.
- [46] Kierstead, H. and Qin, J., Coloring interval graphs with First-Fit. in *Combinatorics of Ordered Sets*, papers from the 4th Oberwolfach Conf., 1991), M. Aigner and R. Wille (eds.), Discrete Math. 144, pp 47-57, 1995.
- [47] Ko, K., Complexity theory of real functions. Progress in Theoretical Computer Science, Birkhäuser Boston, Inc., Boston, MA, 1991.
- [48] Kreitz, C., and Weihrauch, K., Theory of representations, *Theoretical Computer Science*, Volume 38, 1985, Pages 35-53.
- [49] Li, M., and Vitanyi, P., *An Introduction to Kolmogorov Complexity and Its Applications*. Texts and Monographs in Computer Science. Springer-Verlag, 1993.
- [50] Lovasz, L., Saks, M., and Trotter, W., An on-line graph coloring algorithm with sublinear performance ratio, *Discrete Mathematics*, Volume 75, Issues 1–3, May 1989, Pages 319-325.
- [51] Melhorn, K., Polynomial and abstract subrecursive classes. *J. Comput. Syst. Sci.*, 12(2):147–178, 1976.
- [52] Bro Miltersen, P., Subramanian, S., Scott Vitter J., Tamassia R., Complexity models for incremental computation, *Theoretical Computer Science*, Vol. 130 (1994), 203-236.
- [53] Nabutovsky A., and Weinberger S., The fractal nature of Riem/Diff. I. *Geometrica Dedicata*, 101 (2003), 1-54.
- [54] Nies, A., *Computability and Randomness*, volume 51 of Oxford Logic Guides. Oxford University Press, Oxford, 2009.
- [55] Marian B. Pour-El and J. Ian Richards. *Computability in analysis and physics*. Perspectives in Mathematical Logic. Springer-Verlag, Berlin, 1989.
- [56] Robertson N., and Seymour, P., Graph minors II. Algorithmic aspects of tree-width. *Journal of Algorithms*, Vol. 7, pp 309-322, 1986.
- [57] Schröder, M., Extended Admissibility. *Theoretical Computer Science* Vol. 284, 519-538 (2002).
- [58] Sleator, D. and Tarjan, R., Amortized efficiency of list update and paging rules, *Communications of the ACM*, 28 (2) (1985), 202–208.
- [59] Soare, R., Computability theory and differential geometry, *The Bulletin of Symbolic Logic*, 10 (2004), 457–486.

- [60] Turing, A. On Computable Numbers, with an Application to the Entscheidungsproblem, *Proceedings of the London Mathematical Society, Series 2*, Vol. 42 (1936–7), 230–265. Errata appeared in Series 2, 43 (1937), 544–546.
- [61] Wadge, W., “Reducibility and determinateness on the Baire space”. PhD thesis. Univ. of California, Berkeley, 1983.
- [62] Weihrauch, K., *Computable Analysis*, Springer-Verlag, 2000.

VICTORIA UNIVERSITY OF WELLINGTON
Email address: Rod.Downey@msor.vuw.ac.nz

MASSEY UNIVERSITY
Email address: alexander.g.melnikov@gmail.com

NANYANG TECHNOLOGICAL UNIVERSITY
Email address: kmng@ntu.edu.sg