



PAPER

Type-based analysis of logarithmic amortised complexity

Martin Hofmann¹, Lorenz Leutgeb² , David Obwaller³, Georg Moser³ and Florian Zuleger^{2*} 

¹Institut für Informatik, Ludwig-Maximilians-Universität München, Germany, ²Institut für Logic and Computation 192/4, Technische Universität Wien, Austria, and ³Institut für Informatik, Universität Innsbruck, Austria

*Corresponding author. Email: florian.zuleger@tuwien.ac.at

(Received 31 July 2020; revised 7 June 2021; accepted 13 August 2021; first published online 19 October 2021)

Abstract

We introduce a novel amortised resource analysis couched in a type-and-effect system. Our analysis is formulated in terms of the physicist's method of amortised analysis and is potentialbased. The type system makes use of logarithmic potential functions and is the first such system to exhibit *logarithmic amortised complexity*. With our approach, we target the automated analysis of self-adjusting data structures, like splay trees, which so far have only manually been analysed in the literature. In particular, we have implemented a semi-automated prototype, which successfully analyses the zig-zig case of *splaying*, once the type annotations are fixed.

Keywords: Analysis of algorithms; amortised resource analysis; functional programming; self-adjusting data structures; automation

In Memoriam: Martin Hofmann

1. Introduction

Amortised analysis as pioneered by Sleator and Tarjan (1985); Tarjan (1985) is a method for the worst-case cost analysis of data structures. The innovation of amortised analysis lies in considering the cost of a single data structure operation as part of a sequence of data structure operations. The methodology of amortised analysis allows one to assign a low (e.g. constant or logarithmic) amortised cost to a data structure operation even though the worst-case cost of a single operation might be high (e.g. polynomial or worse). The setup of amortised analysis guarantees that for a sequence of data structure operations the worst-case cost is indeed the number of data structure operations times the amortised cost. In this way, amortised analysis provides a methodology for worst-case cost analysis.

Starting with the initial proposal, one of the objectives of amortised analysis has been to conduct a worst-case cost analysis for self-adjusting binary search trees, such as *splay trees* (Sleator and Tarjan, 1985; Tarjan, 1985). These data structures have the behaviour that a single data structure operation might be expensive (i.e. linear in the size of the tree) but the cost is guaranteed to 'average out' in a sequence of data structure operations (i.e. logarithmic in the size of the tree). Amortised analysis has been designed to provide a framework for this kind of reasoning on the cost of data structure operations.

The automated cost analysis of imperative, functional, logic and object-oriented programs as well as of more abstract programming paradigms such as term rewriting systems is an active research topic (Albert et al., 2008, 2011; Alias et al., 2010; Alonso-Blas and Genaim, 2012;

Avanzini et al., 2011, 2015, 2016; Blanc et al., 2010; Flores-Montoya, 2017; Giesl et al., 2017; Gulwani and Zuleger, 2010; Hermenegildo et al., 2012; Hofmann and Moser, 2014, 2015; Moser and Schneckenreither, 2020). Most research has focused on the inference of polynomial bounds on the worst-case cost of the program under analysis. A few papers also target the inference of exponential and logarithmic bounds (Albert et al., 2008; Avanzini et al., 2011; Chatterjee et al., 2017; Kahn and Hoffmann, 2020; Wang et al., 2017; Winkler and Moser, 2020). Some of the cited approaches are able to conduct an automated amortised analysis in the sense of Sleator and Tarjan: The work on type-based cost analysis by Martin Hofmann and his collaborators (Hoffmann, 2011; Hoffmann et al., 2012a, 2017; Hofmann and Jost, 2003; Hofmann and Moser, 2014, 2015; Hoffmann and Shao, 2015a,b; Jost et al., 2009, 2010, 2017), which we discuss in more detail below, directly employs potential functions as proposed in Sleator and Tarjan (1985); Tarjan (1985). For imperative programs, a line of work infers cost bounds from lexicographic ranking functions using arguments that implicitly achieve an amortised analysis, see Sinn et al. (2014, 2015, 2017); Fiedor et al. (2018) and, for details, refer to Sinn et al. (2017)). The connection between ranking functions and amortised analysis has also been discussed in the context of term rewrite systems (Hofmann and Moser, 2014). Proposals that incorporate amortised analysis within the recurrence relations approach to cost analysis have been discussed in Alonso- Blas and Genaim (2012); Flores-Montoya (2017). Still, to the best of our knowledge, none of the cited approaches is able to conduct a worst-case cost analysis for self-adjusting binary search trees such as splay trees. One notable exception is Nipkow (2015) where the correct amortised analysis of splay trees and other data structures is certified in Isabelle/HOL with some tactic support. However, it is not clear at all if the approach can be further automated.

In this article, we take the first step towards the automated analysis of logarithmic amortised cost. We extend the line of work by Martin Hofmann and his collaborators on amortised analysis, where the search for suitable potential functions is encoded as a type-and-effect system. This line of work has led to several successful tools for deriving accurate bounds on the resource usage of functional (Avanzini et al., 2015; Avanzini and Moser, 2016; Hoffmann et al., 2012a; Jost et al., 2009), imperative programs (Hofmann and Rodriguez, 2013; Hoffmann and Sho, 2014), as well as term rewriting systems (Avanzini et al., 2016; Hofmann and Moser, 2014, 2015; Moser and Schneckenreither, 2020). The cited approaches employ a variety of potential functions: While initially confined to inferring linear cost (Hofmann and Jost, 2003), the methods were subsequently extended to cover polynomial (Hofmann and Hofmann, 2010b), multivariate polynomial (Hoffmann et al., 2012a) and also exponential cost (Hofmann and Rodriguez, 2013). We for the first time propose a type system that supports logarithmic potential functions, significantly extending and correcting an earlier note towards this goal (Hofmann and Moser, 2018).

Our analysis is couched in a simple core functional language just sufficiently rich to provide a full definition of our motivating example: *splaying*. We employ a big-step semantics, following similar approaches in the literature. Further, our type system is geared towards runtime as computation cost (i.e. we assign a unit cost to each function call and zero cost to every other program statement). It is straightforward to generalise this type system to other monotone cost models. With respect to non-monotone costs, for example, heap usage, we expect the type system can also be readily adapted, but this is outside the scope of the article.

The type system has been designed with the goal of automation. As in previous work on type-based amortised analysis, the type system infers constraints on unknown coefficients of template potential functions in a syntax-directed way from the program under analysis. Suitable coefficients can then be found automatically by solving the collected constraints with a suitable constraint solver (i.e. an SMT solver that supports the theory of linear arithmetic). The derivation of constraints is straightforward for all syntactic constructs of our programming language. However, our automated analysis also requires *sharing* and *weakening* rules. The latter supports the comparison of different potential functions. As our potential functions are logarithmic, we cannot directly encode the comparison between logarithmic expressions within the theory of linear arithmetic.

Here, we propose several ideas for *linearising* the required comparison of logarithmic expressions. The obtained linear constraints can then be added to the constraint system. Our proposed linearisation makes use of (i) mathematical facts about the logarithm, referred to as *expert knowledge*, (ii) Farkas' Lemma for turning the universally quantified premise of the weakening rule into an existentially quantified statement that can be added to the constraint system and (iii) finally a subtle modification of Schoenmakers potential.

We report on preliminary results for the automated amortised analysis of the splay function. Our implementation semi-automatically verifies the correctness of a type annotation with logarithmic amortised cost for the splay function (more specifically for the zig-zig case of the splay function) using the constraints generated by the type system. We believe that the ideas presented in this article can be extended beyond the case of splay trees to support the analysis of similar self-adjusting data structures such as the ones used in the evaluation of Nipkow (2015). Summarising, we make the following contributions:

- We propose a new class of template potential functions suitable for logarithmic amortised analysis; these potential functions in particular include a variant of Schoenmakers' potential (a key building block for the analysis of the splay function) and logarithmic expressions.
- We present a type-and-effect system for potential-based resource analysis capable of expressing logarithmic amortised costs and prove its soundness.
- We report on a preliminary implementation for the logarithmic amortised analysis of the splay function. With respect to the zig-zig case of the splay function, our prototype is able to automatically check that the amortised cost is bounded by $3 \log(t) + 1$. All former results to this respect required a manual analysis.

Outline. The rest of this article is organised as follows. In Section 2, we review the key concepts underlying type-based amortised analysis and present our ideas for their extension. In Section 3, we introduce a simple core language underlying our reasoning and provide a full definition of playing, our running example. The employed class of potential functions is provided in Section 4, while the type-and-effect system is presented in Section 5. In Section 7, we report on our ideas for implementing the *weakening* rule. Concretely, we see these ideas at work in Section 6, where we employ the established type-and-effect system on the motivating example of *splaying*. In Section 8, we present our implementation and design choices in automation. In Section 9, we present related work and finally, we conclude in Section 10.

2. Setting the Stage

Our analysis is formulated in terms of the physicist's method of amortised analysis in the style of Sleator and Tarjan (1985) and Tarjan (1985). This method assigns a *potential* to data structures of interest and defines the *amortised cost* of an operation as the sum of the actual cost plus the change of the potential through execution of the operation, that is, the central idea of an amortised analysis as formulated by Sleator and Tarjan is to choose a potential function ϕ such that

$$\phi(v) + a_f(v) = c_f(v) + \phi(f(v)),$$

holds for all inputs v to a function f , where a_f, c_f denote the amortised and total cost, respectively, of executing f . Hofmann and Jost (2003); Hoffmann *et al.* (2011, 2012b,a); Hofmann and Moser (2014, 2015) provide a generalisation of this idea to a set of potential functions ϕ, ψ , such that

$$\phi(v) \geq c_f(v) + \psi(f(v)),$$

holds for all inputs v . This allows one to read off an upper bound on the amortised cost of f , that is, we have $a_f(v) \leq \phi(v) - \psi(v)$. We add that the above inequality indeed generalises the original formulation, which can be seen by setting $\phi(v) := a_f(v) + \psi(v)$.

In this article, we present a type-based resource analysis based on the idea of potential functions that can infer logarithmic amortised cost. Following previous work by Hofmann et al., we tackle two key problems to achieve a semi-automated logarithmic amortised analysis: (1) Automation is achieved by a *type-and-effect system* that uses *template potential functions*, that is, functions of a fixed shape with indeterminate coefficients. Here, the key challenge is to identify templates that are suitable for logarithmic analysis and that are closed under the basic operations of the considered programming language. (2) In addition to the actual amortised analysis with costs, we employ *cost-free* analysis as a subroutine, setting the amortised a_f and actual costs c_f of all functions f to zero. This enables a *size analysis* of sorts, because the inequality $\phi(v) \geq \psi(f(v))$ bounds the size of the potential $\psi(f(v))$ in terms of the potential $\phi(v)$. The size analysis we conduct allows lifting the analysis of a subprogram to a larger context, which is crucial for achieving a *compositional analysis*. We overview these two aspects in the sequel of the section.

2.1 Type-and-effect system

To set the scene, we briefly review amortised analysis formulated as a type-and-effect system up to and including the multivariate polynomial analysis, cf. Jost et al. (2010); Hoffmann and Hofmann (2010b,a); Hoffmann et al. (2011, 2012a); Hofmann and Moser (2014, 2015); Hoffmann et al. (2017); Jost et al. (2017).

Polynomial Amortised Analysis. Suppose that we have types $\alpha, \beta, \gamma, \dots$ representing sets of values. We write $\llbracket \alpha \rrbracket$ for the set of values represented by type α . Types may be constructed from base types such as Booleans and integers, denoted by *Base*, and by type formers such as list, tree, product and sum. For each type α , we define a (possibly infinite) set of *basic potential functions* $\mathcal{B}(\alpha): \llbracket \alpha \rrbracket \rightarrow \mathbb{R}_0^+$. Thus, if $p \in \mathcal{B}(\alpha)$ and $v \in \llbracket \alpha \rrbracket$, then $p(v) \in \mathbb{R}_0^+$. An *annotated type* is a pair of a type α and a function $Q: \mathcal{B}(\alpha) \rightarrow \mathbb{R}_0^+$ providing a coefficient for each *basic potential function*. The function Q must be zero on all but finitely many basic potential functions. For each annotated type $\alpha|Q$, the *potential function* $\phi_Q: \llbracket \alpha \rrbracket \rightarrow \mathbb{R}_0^+$ is then given by

$$\phi_Q(v) := \sum_{p \in \mathcal{B}(\alpha)} Q(p) \cdot p(v).$$

By introducing product types, one can regard functions with several arguments as unary functions, which allows for technically smooth formalisations, cf. Hoffmann and Hofmann (2010b,a); Hoffmann (2011); the analyses in the cited papers are called *univariate* as the set of basic potential functions $\mathcal{B}(\alpha)$ of a product type α is given directly. In the later *multivariate* versions of automated amortised analysis (Hoffmann et al., 2011, 2012a; Hofmann and Moser, 2015) one takes a more fine-grained approach to products. Namely, one then sets (for arbitrary n)

$$\begin{aligned} \mathcal{B}(\alpha_1 \times \dots \times \alpha_n) &:= \mathcal{B}(\alpha_1) \times \dots \times \mathcal{B}(\alpha_n) \\ ((p_1, \dots, p_n)((v_1, \dots, v_n)) &:= \prod_{i=1}^n p_i(v_i). \end{aligned}$$

Thus, the basic potential function for a product type is obtained as the multiplication of the basic potential functions of its constituents.¹

Automation. The idea behind this setup is that the basic potential functions $\mathcal{B}(\alpha)$ are suitably chosen and fixed by the analysis designer; the coefficients $Q(p)$ for $p \in \mathcal{B}(\alpha)$, however, are left indeterminate and will (automatically) be fixed during the analysis. For this, constraints over

¹Suppose that for each type α there exists a distinguished element $u \in \mathcal{B}(\alpha)$ with $u(a) = 1$ for all $a \in \llbracket \alpha \rrbracket$. Then, the multivariate product types contain all (linear combinations) of the basic potential functions, extending earlier univariate definitions of product types.

the unknown coefficients are collected in a syntax-directed way from the function under analysis and then solved by a suitable constraint solver. The type-and-effect system formalises this collection of constraints as typing rules, where for each construct of the considered programming language a typing rule is given that corresponds to constraints over the coefficients of the annotated types. Expressing the quest for suitable type annotations as a type-and-effect system allows one to compose typing judgements in a syntax-oriented way without the need for fixing additional intermediate results, which is often required by competing approaches. This syntax-directed approach to amortised analysis has been demonstrated to work well for data types like lists or trees whose basic potential functions are polynomials over the length of a list resp. the number of nodes of a tree. One of the reasons why this works well is, for example, that functional programming languages typically include dedicated syntax for list construction and that polynomials are closed under addition by one (i.e. if $p(n)$ is a polynomial, so is $p(n + 1)$), supporting the formulation of a suitable typing rule for list construction, cf. Hoffmann and Hofmann (2010b,a); Hoffmann (2011); Hoffmann et al. (2011, 2012a). The syntax-directed approach has been shown to generalise from lists and trees to general inductive data types, cf. Hofmann and Moser (2014, 2015); Moser and Schneckenreither (2018).

Logarithmic Amortised Analysis. We now motivate the design choices of our type-and-effect system. The main objective of our approach is the automated analysis of data structures such as splay trees, which have *logarithmic* amortised cost. The amortised analysis of splay trees is tricky and requires choosing an adequate potential function: our work makes use of a variant of Schoenmakers’ potential, $\text{rk}(t)$ for a tree t , cf. Schoenmakers (1993); Nipkow (2015), defined inductively by

$$\begin{aligned} \text{rk}(\text{leaf}) &:= 1, \\ \text{rk}(l, d, r) &:= \text{rk}(l) + \log(|l|) + \log(|r|) + \text{rk}(r), \end{aligned}$$

where l, r are the left resp. right child of the tree (l, d, r) , $|t|$ denotes the number of leaves of a tree t , and d is some data element that is ignored by the potential function. Besides Schoenmakers’ potential we need to add further basic potential functions to our analysis. This is motivated as follows: Similar to the polynomial amortised analysis discussed above, we want that the basic potential functions can express the construction of a tree, for example, let us consider the function

$$f(x, d, y) := (x, d, y),$$

which constructs the tree (x, d, y) from some trees x, y and some data element d , and let us assume a constant cost $c_f(x, y) = 1$ for the function f . A type annotation for f is given by

$$\underbrace{\text{rk}(x) + \log(|x|) + \text{rk}(y) + \log(|y|) + 1}_{\phi(x,y)} \geq c_f(x, y) + \underbrace{\text{rk}(f(x, d, y))}_{\psi(f(x,y))},$$

that is, the potential $\phi(x, y)$ suffices to pay for the cost c_f of executing f and the potential of the result $\psi(f(x, y))$ (the correctness of this annotation can be established directly from the definition of Schoenmakers’ potential). As mentioned above, the logarithmic expressions in $\phi(x, y)$, that is, $\log(|x|) + \log(|y|) + 1$, specify the *amortised costs* of the operation.

We see that to express the potential $\phi(x, y)$ we also need the basic potential functions $\log(|t|)$ for a tree t . In fact, we will choose the slightly richer set of basic potential functions

$$p_{(a,b)}(t) = \log(a \cdot |t| + b),$$

where $a, b \in \mathbb{N}$ and t is a tree. We note that by setting $a = 0$ and $b = 2$ this choice allows us to represent the constant function u with $u(t) = 1$ for all trees t . We further note that this choice of potential functions is sufficiently rich to express that $p_{(a,b)}(t) = p_{(a,b+a)}(s)$ for trees s, t with $|t| =$

$|s| + 1$, which is needed for precisely expressing the change of potential when a tree is extended by one node. Further, we define basic potential functions for products of trees by setting

$$P_{((a_1, \dots, a_n), b)}((t_1, \dots, t_n)) = \log(a_1 \cdot |t_1| + \dots + a_n \cdot |t_n| + b)$$

where $(a_1, \dots, a_n), b \in \mathbb{N}$ and (t_1, \dots, t_n) is a tuple of trees. This is sufficiently rich to state the equality $P_{(a_0, (a_1, \dots, a_n), b)}(x_1, (x_1, \dots, x_n)) = P_{(a_0 + (a_1, \dots, a_n), b)}((x_1, \dots, x_n))$, which supports the formulation of a *sharing* rule, which in turn is needed for supporting the *let*-construct in functional programming; cf. Hoffmann et al. (2011, 2012a); Hofmann and Moser (2015) for a more detailed exposition on the sharing rule and the *let*-construct.

2.2 Cost-free semantics

Polynomial Amortised Analysis. We begin by reviewing the cost-free semantics underlying previous work (Hoffmann, 2011; Hoffmann et al., 2011, 2012a; Hoffmann and Hofmann, 2010a) on polynomial amortised analysis. Assume that we want to analyse the composed function call $g(f(\vec{x}), \vec{z})$ using already established analysis results for $f(\vec{x})$ and $g(y, \vec{z})$. Suppose we have already established that for all \vec{x}, y, \vec{z} we have:

$$\phi_0(\vec{x}) \geq c_f(\vec{x}) + \beta(f(\vec{x})) \tag{1}$$

$$\phi_i(\vec{x}) \geq \phi'_i(f(\vec{x})) \quad \text{for all } i \ (0 < i \leq n) \tag{2}$$

$$\beta(y) + \gamma(\vec{z}) + \sum_{i=1}^n \phi'_i(y)\phi''_i(\vec{z}) \geq c_g(y, \vec{z}) + \psi(g(y, \vec{z})), \tag{3}$$

where as in the multivariate case above, n is arbitrary and equations (1) and (3) assume cost, while equation (2) is *cost-free*. Then, we can conclude for all \vec{x}, \vec{z} that

$$\underbrace{\phi_0(\vec{x}) + \gamma(\vec{z}) + \sum_{i=1}^n \phi_i(\vec{x})\phi''_i(\vec{z})}_{\phi(\vec{x}, \vec{z})} \geq c_f(\vec{x}) + c_g(f(\vec{x}), \vec{z}) + \psi(g(f(\vec{x}), \vec{z})),$$

guaranteeing that the potential $\phi(\vec{x}, \vec{z})$ suffices to pay for the cost $c_f(\vec{x})$ of computing $f(\vec{x})$, the cost $c_g(f(\vec{x}), \vec{z})$ of computing $g(f(\vec{x}), \vec{z})$ and the potential $\psi(g(f(\vec{x}), \vec{z}))$ of the result $g(f(\vec{x}), \vec{z})$. We note that the correctness of this inference hinges on the fact that we can multiply equation (2) with $\phi''_i(\vec{z})$ for $i = 1 \dots n$, using the monotonicity of the multiplication operation (note that potential functions are non-negative). We highlight that the multiplication argument works well with cost-free semantics and enables lifting the resource analysis of $f(\vec{x})$ and $g(y, \vec{z})$ to the composed function call $g(f(\vec{x}), \vec{z})$.

Remark. We point out that the above exposition of cost-free semantics in the context of polynomial amortised analysis differs from the motivation given in the literature (Hoffmann, 2011; Hoffmann et al., 2011, 2012a; Hoffmann and Hofmann, 2010a), where cost-free semantics are motivated by the quest for *resource polymorphism*, which is the problem of computing (a representation of) all polynomial potential functions (up to a fixed maximal degree) for the program under analysis; this problem has been deemed of importance for the handling of non-tail-recursive programs. We add that for the amortised cost analysis of inductively generated data types, the cost-free semantics proved necessary even for handling basic data structure manipulations (Hofmann and Moser, 2014, 2015; Moser and Schneckeneither, 2020). In our view, cost-free semantics incorporate a *size analysis* of sorts. We observe that equation (2) states that the potential of the result of the evaluation of $f(\vec{x})$ is bounded by the potential of the function arguments \vec{x} , without accounting for the costs of this evaluation. Thus, for suitably chosen potential functions ϕ_i, ϕ'_i can act as *norms* and capture the size of the result of the evaluation $f(\vec{x})$ in relation to the

size of the argument. As stated above, a separated cost and size analysis enables a compositional analysis, an insight that we also exploit for logarithmic amortised analysis.

Logarithmic Amortised Analysis. Similar to the polynomial case, we want to analyse the composed function call $g(f(\vec{x}), \vec{z})$ using already established analysis results for $f(\vec{x})$ and $g(y, \vec{z})$. However, now we extend the class of potential functions to sublinear functions. Assume that we have already established that

$$\phi_0(\vec{x}) \geq c_f(\vec{x}) + \beta(f(\vec{x})) \tag{4}$$

$$\log(\phi_i(\vec{x})) \geq \log(\phi_i'(f(\vec{x}))) \quad \text{for all } i \ (0 < i \leq n) \tag{5}$$

$$\beta(y) + \gamma(\vec{z}) + \sum_{i=1}^n \log(\phi_i'(y) + \phi_i''(\vec{z})) \geq c_g(y, \vec{z}) + \psi(g(y, \vec{z})), \tag{6}$$

where equations (4) and (6) assume cost, while equation (5) is *cost-free*. Equations (4) and (5) represent the result of an analysis of $f(\vec{x})$ (note that these equations do not contain the parameters \vec{z} , which will however be needed for the analysis of $g(f(\vec{x}), \vec{z})$), and equation (6) the result of an analysis of $g(y, \vec{z})$. Then, we can conclude for all \vec{x}, y, \vec{z} that

$$\underbrace{\phi_0(\vec{x}) + \gamma(\vec{z}) + \sum_{i=1}^n \log(\phi_i(\vec{x}) + \phi_i''(\vec{z}))}_{\phi(\vec{x}, \vec{z})} \geq c_f(\vec{x}) + c_g(f(\vec{x}), \vec{z}) + \psi(g(f(\vec{x}), \vec{z})),$$

guaranteeing that the potential $\phi(\vec{x}, \vec{z})$ suffices to pay for the cost $c_f(\vec{x})$ of computing $f(\vec{x})$, the cost $c_g(f(\vec{x}), \vec{z})$ of computing $g(f(\vec{x}), \vec{z})$ and the potential $\psi(g(f(\vec{x}), \vec{z}))$ of the result $g(f(\vec{x}), \vec{z})$. Here, we crucially use monotonicity of the logarithm function, as formalised in Lemma 13. This reasoning allows us to lift isolated analyses of the functions $f(\vec{x})$ and $g(y, \vec{z})$ to the composed function call $g(f(\vec{x}), \vec{z})$; this is what is required for a compositional analysis!

Example. We now illustrate the compositional reasoning on an example. We reconsider the function $f(x, d, y) := (x, d, y)$, which takes two trees x, y and some data element d and returns the tree (x, d, y) . Assume that we already have established that

$$\psi(x) + \psi(y) + 1 \geq c_f(x, y) + \text{rk}(f(x, d, y)) \tag{7}$$

$$\log(|x| + |y|) \geq \log(|f(x, d, y)|), \tag{8}$$

where $\psi(u) = \text{rk}(u) + \log(|u|)$, $c_f(x, y) = 1$, and d is an arbitrary data element, which is not relevant for the cost analysis of f . We now want to analyse the composed function $h(x, a, y, b, z) := f(f(x, a, y), b, z)$. We will use the above reasoning, instantiating equations (4) and (5) with equations (7) and (8) for the inner function call $f(x, a, y)$, and equation (6) with the sum of equations (7) and (8) for the outer function call $f(u, b, z)$. As argued above, we can then conclude for all x, y, z that

$$\begin{aligned} \psi(x) + \psi(y) + \psi(z) + \log(|x| + |y|) + \log(|x| + |y| + |z|) + 2 &\geq \\ &\geq c_f(x, a, y) + c_f(f(x, a, y), b, z) + \psi(f(f(x, a, y), z)), \end{aligned}$$

is a valid type annotation for $h(x, a, y, b, z) := f(f(x, a, y), b, z)$; we have used equation (8) twice in this derivation, once as $\log(|x| + |y|) \geq \log(|f(x, a, y)|)$ and once lifted as $\log(|x| + |y| + |z|) \geq \log(|f(x, a, y)| + |z|)$. Kindly note that the above example appears in similar form as part of the analysis of the `splay` function described in Section 6.

3. Motivating Example

In this section, we introduce the syntax of a suitably defined core (first-order) programming language to be used in the following. Furthermore, we recall the definition of *splaying*, following the

presentation in Nipkow (2015). Splaying constitutes the motivating example for the type-based logarithmic amortised resource analysis presented in this article.

To make the presentation more succinct, we assume only the following types: Booleans $\text{Bool} = \{\text{true}, \text{false}\}$, an abstract base type Base (abbrev. B), product types and binary trees Tree (abbrev. T), whose internal nodes are labelled with elements $b : \text{Base}$. We use lower-case Greek letters for the denotation of types. Elements $t : \text{Tree}$ are defined by the following grammar which fixes notation.

$$t ::= \text{leaf} \mid (t_1, b, t_2).$$

The size of a tree is the number of leaves: $|\text{leaf}| := 1, |(t, a, u)| := |t| + |u|$.

Expressions are defined as follows and given in *let normal form* to simplify the presentation of the semantics and typing rules. To ease the readability, we make use of some mild syntactic sugaring in the presentation of actual code.

Definition 1.

$$\begin{aligned} o &::= < \mid > \mid = \\ e &::= f \ x_1 \ \dots \ x_n \\ & \mid \text{true} \mid \text{false} \mid e_1 \circ e_2 \quad \mid \text{if } x \text{ then } e_1 \text{ else } e_2 \\ & \mid (x_1, x_2, x_3) \mid \text{leaf} \quad \mid \text{match } x \text{ with } \mid \text{leaf} \rightarrow e_1 \mid (x_1, x_2, x_3) \rightarrow e_2 \\ & \mid \text{let } x = e_1 \text{ in } e_2 \quad \mid x \end{aligned}$$

We skip the standard definition of integer constants $n \in \mathbb{Z}$ as well as variable declarations, cf. Pierce (2002). Furthermore, we omit binary operators and only define essential comparisons. For our analysis, these are unimportant, as long as we assume that no actual costs are emitted.

A *typing context* is a mapping from variables \mathcal{V} to types. Type contexts are denoted by upper-case Greek letters. A program P consists of a signature \mathcal{F} together with a set of function definitions of the form $f \ x_1 \ \dots \ x_n = e$, where the x_i are variables and e an expression. A *substitution* or (*environment*) σ is a mapping from variables to values that respects types. Substitutions are denoted as sets of assignments: $\sigma = \{x_1 \mapsto t_1, \dots, x_n \mapsto t_n\}$. We write $\text{dom}(\sigma)$ ($\text{rg}(\sigma)$) to denote the domain (range) of σ . Let σ, τ be substitutions such that $\text{dom}(\sigma) \cap \text{dom}(\tau) = \emptyset$. Then we denote the (disjoint) union of σ and τ as $\sigma \uplus \tau$. We employ a simple cost-sensitive big-step semantics based on eager evaluation, whose rules are given in Figure 1. The judgement $\sigma \stackrel{\ell}{\vdash} e \Rightarrow v$ means that under environment σ , expression e is evaluated to value v in exactly ℓ steps. Here only rule applications emit (unit) costs. If we do not take costs into account, we simply write $\sigma \vdash e \Rightarrow v$.

Splay Trees have been introduced in Sleator and Tarjan (1985); Tarjan (1985) as self-adjusting binary search trees with strictly increasing in-order traversal. There is no explicit balancing condition. All operations rely on a tree rotating operation dubbed *splaying*; `splay a t` is performed by rotating element a to the root of tree t while keeping in-order traversal intact. If a is not contained in t , then the last element found before `leaf` is rotated to the tree. The complete definition is given in Figure 2. Based on splaying, searching is performed by splaying with the sought element and comparing to the root of the result. Similarly, the definition of insertion and deletion depends on splaying. As an example, the definition of `insert` and `delete` is given in Figures 3 and 4, respectively. See also Nipkow (2015) for full algorithmic, formally verified, descriptions.

All basic operations can be performed in $O(\log n)$ amortised runtime. The logarithmic amortised complexity is crucially achieved by local rotations of subtrees in the definition of `splay`. Amortised cost analysis of splaying has been provided, for example, in Sleator and Tarjan (1985); Schoenmakers (1993); Nipkow (2015); Okasaki (1999), among others. Below, we follow Nipkow’s approach, where the actual cost of splaying is measured by counting the number of calls to `splay`: $\text{B} \times \text{T} \rightarrow \text{T}$.

$$\begin{array}{c}
 \frac{}{\sigma \mid^0 \text{false} \Rightarrow \text{false}} \qquad \frac{}{\sigma \mid^0 \text{true} \Rightarrow \text{true}} \qquad \frac{}{\sigma \mid^0 \text{leaf} \Rightarrow \text{leaf}} \\
 \\
 \frac{x_1 \sigma = t \quad x_2 \sigma = b \quad x_3 \sigma = u}{\sigma \mid^0 (x_1, x_2, x_3) \Rightarrow (t, b, u)} \qquad \frac{x \sigma = v}{\sigma \mid^0 x \Rightarrow v} \qquad \frac{b \text{ is value of } x_1 \sigma \circ x_2 \sigma}{\sigma \mid^0 x_1 \circ x_2 \Rightarrow b} \\
 \\
 \frac{f \ y_1 \ \dots \ y_k = e \in \mathsf{P} \quad \sigma' \mid^\ell e \Rightarrow v}{\sigma \mid^{\ell+1} f \ x_1 \ \dots \ x_k \Rightarrow v} \qquad \frac{\sigma \mid^{\ell_1} e_1 \Rightarrow w \quad \sigma[x \mapsto w] \mid^{\ell_2} e_2 \Rightarrow v}{\sigma \mid^{\ell_1 + \ell_2} \text{let } x = e_1 \text{ in } e_2 \Rightarrow v} \\
 \\
 \frac{x \sigma = \text{leaf} \quad \sigma \mid^\ell e_1 \Rightarrow v}{\sigma \mid^\ell \text{match } x \text{ with } \mid \text{leaf} \rightarrow e_1 \quad \Rightarrow v} \qquad \frac{x \sigma = \text{false} \quad \sigma \mid^\ell e_2 \Rightarrow v}{\sigma \mid^\ell \text{if } x \text{ then } e_1 \text{ else } e_2 \Rightarrow v} \\
 \qquad \qquad \qquad \mid (x_0, x_1, x_2) \rightarrow e_2 \\
 \\
 \frac{x \sigma = (t, a, u) \quad \sigma' \mid^\ell e_2 \Rightarrow v}{\sigma \mid^\ell \text{match } x \text{ with } \mid \text{leaf} \rightarrow e_1 \quad \Rightarrow v} \qquad \frac{x \sigma = \text{true} \quad \sigma \mid^\ell e_1 \Rightarrow v}{\sigma \mid^\ell \text{if } x \text{ then } e_1 \text{ else } e_2 \Rightarrow v} \\
 \qquad \qquad \qquad \mid (x_0, x_1, x_2) \rightarrow e_2
 \end{array}$$

Here $\sigma[x \mapsto w]$ denotes the update of the environment σ such that $\sigma[x \mapsto w](x) = w$ and the value of all other variables remains unchanged. Furthermore, in the second `match` rule, we set $\sigma' := \sigma \uplus \{x_0 \mapsto t, x_1 \mapsto a, x_2 \mapsto u\}$ and for function application we set $\{y_1 \mapsto x_1 \sigma, \dots, y_k \mapsto x_k \sigma\}$.

Figure 1. Big-step semantics.

4. Resource Functions

In this section, we detail the basic potential functions employed and clarify the definition of potentials used.

Only trees are assigned non-zero potential. This is not a severe restriction as potentials for basic data types would only become essential if the construction of such types would emit actual costs. This is not the case in our context. Moreover, note that lists can be conceived as trees of particular shape. The potential $\Phi(t)$ of a tree t is given as a non-negative linear combination of basic functions, which essentially amount to ‘sums of logs’, cf. Schoenmakers (1993). It suffices to specify the basic functions for the type of trees T . As already mentioned in Section 2, the *rank* $\text{rk}(t)$ of a tree is defined as follows:

$$\begin{aligned}
 \text{rk}(\text{leaf}) &:= 1 \\
 \text{rk}((t, a, u)) &:= \text{rk}(t) + \log'(|t|) + \log'(|u|) + \text{rk}(u) .
 \end{aligned}$$

We set $\log'(n) := \log_2(\max\{n, 1\})$, that is, the (binary) logarithm function is defined for all numbers. This is merely a technicality, introduced to ease the presentation as it simplifies the statement of subsequent definitions. In the following, we will denote the modified logarithmic function, simply as \log . Furthermore, recall that $|t|$ denotes the number of leaves in tree t . The definition of ‘rank’ is inspired by the definition of potential in Schoenmakers (1993); Nipkow (2015), but subtly changed to suit it to our context.

Definition 2. *The basic potential functions of Tree , denoted \mathcal{B} , are*

- $\lambda t. \text{rk}(t)$, and
- $p_{(a,b)} := \lambda t. \log(a \cdot |t| + b)$, where a, b are natural numbers.

```
splay a t = match t with
| leaf      -> leaf
| (cl, c, cr) -> if a == c
  then (cl, c, cr)
  else if a < c
    then match cl with
      | leaf      -> (leaf, c, cr)
      | (bl, b, br) -> if a == b
        then (bl, a, (br, c, cr))
        else if a < b
          then match bl with
            | leaf -> (leaf, b, (br, c, cr))
            | bl  -> match splay a bl with
              | leaf      -> leaf
              | (al, a1, ar) -> (al, a1, (ar, b, (br, c, cr)))
          else match br with
            | leaf -> (bl, b, (leaf, c, cr))
            | br  -> match splay a br with
              | leaf      -> leaf
              | (al, a1, ar) -> ((bl, b, al), a1, (ar, c, cr))
    else match cr with
      | leaf      -> (cl, c, leaf)
      | (bl, b, br) -> if a == b
        then ((cl, c, bl), a, br)
        else if a < b
          then match bl with
            | leaf -> ((cl, c, leaf), b, br)
            | bl  -> match splay a bl with
              | leaf      -> leaf
              | (al, a1, ar) -> ((cl, c, al), a1, (ar, b, br))
          else match br with
            | leaf -> ((cl, c, bl), b, leaf)
            | br  -> match splay a br with
              | leaf      -> leaf
              | (al, a1, ar) -> (((cl, c, bl), b, al), a1, ar)
```

Figure 2. Function splay.

```
insert a t = match t with
| leaf -> (leaf, a, leaf)
| t    -> match splay a t with
  | leaf      -> leaf
  | (l, b, r) -> if a == b
    then (l, a, r)
    else if a < b
      then (l, a, (leaf, b, r))
      else ((l, b, leaf), a, r)
```

Figure 3. Function insert.

Note that the constant function 1 is representable: $1 = \lambda t. \log(0 \cdot |t| + 2) = p_{(0,2)}$.

Following the recipe of the high-level description in Section 2, potentials or more generally resource functions become definable as linear combinations of basic potential functions.

Definition 3. A resource function $r: \llbracket \mathbb{T} \rrbracket \rightarrow \mathbb{R}_0^+$ is a non-negative linear combination of basic potential functions, that is,

$$r(t) := \sum_{i \in I} q_i \cdot p_i(t),$$

where $q_i \in \mathbb{R}_0^+$, $p_i \in \mathcal{B}$ and $I := \{*\} \cup (\mathbb{N} \times \mathbb{N})$. The set of resource functions is denoted \mathcal{R} .

```

delete a t = match splay a t with
| leaf      -> leaf
| (l, b, r) -> if a == b
    then match l with
        | leaf -> r
        | l   -> match splay_max l with
            | leaf      -> leaf
            | (ll, m, _) -> (ll, m, r)
    else (l, b, r)

splay_max t = match t with
| leaf      -> leaf
| (l, b, r) -> match r with
| leaf      -> (l, b, leaf)
| (rl, c, rr) -> match rr with
| leaf -> ((l, b, rl), c, leaf)
| rr   -> match splay_max rr with
| leaf      -> leaf
| (rrll, x, xa) -> (((l, b, rl), c, rrll), x, xa)

```

Figure 4. Functions delete and splay_max.

A resource annotation over \mathbb{T} , or simply annotation, is a sequence $Q = [q_*] \cup [(q_{(a,b)})_{a,b \in \mathbb{N}}]$ with $q_*, q_{(a,b)} \in \mathbb{Q}_0^+$ with all but finitely many of the coefficients $q_*, q_{(a,b)}$ equal to 0. It represents a (finite) linear combination of basic potential functions, that is, a resource function. The empty annotation, that is, the annotation where all coefficients are set to zero, is denoted as \emptyset .

Remark 4. We use the convention that the sequence elements of resource annotations are denoted by the lower-case letter of the annotation, potentially with corresponding sub- or superscripts.

Definition 5. The potential of a tree t with respect to an annotation $Q = [q_*] \cup [(q_{(a,b)})_{a,b \in \mathbb{N}}]$, is defined as follows:

$$\Phi(t|Q) := q_* \cdot \text{rk}(t) + \sum_{a,b \in \mathbb{N}} q_{(a,b)} \cdot p_{(a,b)}(t),$$

Recall that $p_{(a,b)} = \log(a \cdot |t| + b)$ and that rk is the rank function, defined above.

Example 1. Let t be a tree, then its potential could be defined as follows: $\text{rk}(t) + 3 \cdot \log(|t|) + 1$. With respect to the above definition, this potential becomes representable by setting $q_* := 1, q_{(1,0)} := 3, q_{(0,2)} := 1$. Thus, $\Phi(t|Q) = \text{rk}(t) + 3 \cdot \log(|t|) + 1$. □

We emphasise that the linear combination defined above is not independent. Consider, for example, $\log(2|t| + 2) = \log(|t| + 1) + 1$.

Analysis of Products of Trees. We now lift the basic potential functions $p_{(a,b)}$ of a single tree to products of trees. As discussed in Section 2, we define the potential functions $p_{((a_1, \dots, a_m), b)}$ for a sequence of m trees (t_1, \dots, t_m) , by setting:

$$p_{((a_1, \dots, a_m), b)}((t_1, \dots, t_m)) := \log(a_1 \cdot |t_1| + \dots + a_m \cdot |t_m| + b)$$

where $(a_1, \dots, a_m), b \in \mathbb{N}$. Equipped with this definition, we generalise annotations to sequences of trees. An annotation for a sequence of length m is a sequence $Q = [q_1, \dots, q_m] \cup [(q_{((a_m, \dots, a_n), b)})_{a_i, b \in \mathbb{N}}]$, again vanishing almost everywhere. Note that an annotation of length 1

is simply an annotation as defined above, where the coefficient q_1 is set equal to the coefficient q_* . Based on this, the potential of a sequence of trees (t_1, \dots, t_m) is defined as follows:

Definition 6. Let (t_1, \dots, t_m) be trees and let $Q = [(q_1, \dots, q_m)] \cup [(q_{((a_1, \dots, a_m), b)})_{a_i, b \in \mathbb{N}}]$ be an annotation of length m as above. We define

$$\Phi((t_1, \dots, t_m)|Q) := \sum_{i=1}^m q_i \cdot \text{rk}(t_i) + \sum_{(a_1, \dots, a_m), b \in \mathbb{N}} q_{((a_1, \dots, a_m), b)} \cdot p_{((a_1, \dots, a_m), b)}((t_1, \dots, t_m)),$$

where $p_{((a_1, \dots, a_m), b)}((t_1, \dots, t_m)) := \log(a_1 \cdot |t_1| + \dots + a_m \cdot |t_m| + b)$ as defined above. We use $|Q|$ to denote the length of Q .

Note that for an empty sequence of trees, we have $\Phi(\epsilon|Q) := \sum_{b \in \mathbb{N}} q_b \log(b)$. Note that the rank function $\text{rk}(t)$ amounts to the sum of the logarithms of the size of subtrees of t . In particular, if the tree t simplifies to a list of length n , then $\text{rk}(t) = (n + 1) + \sum_{i=1}^n \log(i)$. Moreover, as $\sum_{i=1}^n \log(i) \in \Theta(n \log n)$, the above-defined potential functions are sufficiently rich to express linear combinations of sub- and super-linear functions.

Let σ denote a substitution, let Γ denote a typing context and let $x_1 : T, \dots, x_m : T$ denote all tree types in Γ . A resource annotation for Γ or simply annotation is an annotation for the sequence of trees $x_1\sigma, \dots, x_m\sigma$. We define the potential of $\Gamma : Q$ with respect to σ as $\Phi(\sigma; \Gamma|Q) := \Phi(x_1\sigma, \dots, x_m\sigma|Q)$.

Definition 7. An annotated signature $\overline{\mathcal{F}}$ is a mapping from functions f to sets of pairs consisting of the annotation type for the arguments of f , $\alpha|_1 \times \dots \times \alpha_n|Q$ and the annotation type $\beta|Q'$ for the result:

$$\overline{\mathcal{F}}(f) := \left\{ \alpha|_1 \rightarrow 1| \times \dots \times \alpha_n|Q\beta Q' : f \text{ takes } n \text{ arguments of which } m \text{ are trees, } |Q| = m, |Q'| = \right\}.$$

Note that $m \leq n$ by definition.

We confuse the signature and the annotated signature and denote the latter simply as \mathcal{F} . Instead of $\alpha|_1 \rightarrow 1| \times \dots \times \alpha_n|Q\beta Q' \in \overline{\mathcal{F}}(f)$, we typically write $f : \alpha_1| \times \dots \times \alpha_n|Q\beta Q'$. As our analysis makes use of a cost-free semantics any function symbol is possibly equipped with a cost-free signature, independent of \mathcal{F} . The cost-free signature is denoted as \mathcal{F}^{cf} .

Example 2. Consider the function *splay*: $B \times T \rightarrow T$. The induced annotated signature is given as $B \times T|Q \rightarrow T|Q'$, where $Q := [q_*] \cup [(q_{(a,b)})_{a,b \in \mathbb{N}}]$ and $Q' := [q'_*] \cup [(q'_{(a,b)})_{a,b \in \mathbb{N}}]$. The logarithmic amortised cost of splaying is then expressible through the following setting: $q_* := 1, q_{(1,0)} = 3, q_{(0,2)} = 1, q'_* := 1$. All other coefficients are zero.

This amounts to a potential of the arguments $\text{rk}(t) + 3 \log(|t|) + 1$, while for the result we consider only its rank, that is, the annotation expresses $3 \log(|t|) + 1$ as the logarithmic cost of splaying. The correctness of the induced logarithmic amortised costs for the zig-zig case of splaying is verified in Section 6 and is also automatically verified by our prototype. \square

Suppose $\Phi((t_1, \dots, t_n), u_1, u_2|Q)$ denotes an annotated sequence of length $n + 2$. Suppose further $u_1 = u_2 =: u$ and we want to share the value u , that is, the corresponding function argument appears multiple times in the body of the function definition. Then we make use of the operator $\Upsilon(Q)$ that adapts the potential suitably. The operator is also called *sharing operator* (in analogy to Hoffmann et al., 2012a, Lemma 6.6).

Lemma 8. Let $(t_1, \dots, t_n), u_1, u_2$ denote a sequence of trees of length $n + 2$ with annotation Q . Then there exists a resource annotation $\Upsilon(Q)$ such that $\Phi((t_1, \dots, t_n), u_1, u_2|Q) = \Phi((t_1, \dots, t_n), u|\Upsilon(Q))$, if $u_1 = u_2 = u$.

Proof. Wlog. we assume $n = 0$. Thus, let $Q = [q_1, q_2] \cup [(q_{(a_1, a_2, b)})_{a_1, a_2, b \in \mathbb{N}}]$. By definition

$$\Phi(u_1, u_2 | Q) = q_1 \cdot \text{rk}(u_1) + q_2 \cdot \text{rk}(u_2) + \sum_{a_1, a_2, b \in \mathbb{N}} q_{(a_1, a_2, b)} \cdot p_{(a_1, a_2, b)}(u_1, u_2),$$

where $p_{(a_1, a_2, b)}(u_1, u_2) = \log(a_1 \cdot |u_1| + a_2 \cdot |u_2| + b)$. By assumption $u = u_1 = u_2$. Thus, we obtain

$$\begin{aligned} \Phi(u, u | Q) &= q_1 \cdot \text{rk}(u) + q_2 \cdot \text{rk}(u) + \sum_{a_1, a_2, b \in \mathbb{N}} q_{(a_1, a_2, b)} \cdot p_{(a_1, a_2, b)}(u, u) \\ &= (q_1 + q_2) \text{rk}(u) + \sum_{a_1 + a_2, b \in \mathbb{N}} q_{(a_1 + a_2, b)} \cdot p_{(a_1 + a_2, b)}(u) \\ &= \Phi(u | \Upsilon(Q)), \end{aligned}$$

for suitable defined annotation $\Upsilon(Q)$, whose definition can be directly read off from the above constraints. □

We emphasise that the definability of the sharing annotation $\Upsilon(Q)$ is based on the fact that the basic potential functions $p_{((a_1, \dots, a_m), b)}$ have been carefully chosen so that

$$p_{(a_0, a_1, a_2, \dots, a_m, b)}(x_1, (x_1, \dots, x_m)) = p_{(a_0 + a_1, a_2, \dots, a_m, b)}(x_1, x_2, \dots, x_m),$$

holds, cf. Section 2.

Remark 9. We observe that the proof-theoretic analogue of the sharing operation constitutes in a contraction rule, if the type system is conceived as a proof system.

Let $Q = [q_*] \cup [(q_{(a,b)})_{a,b \in \mathbb{N}}]$ be an annotation and let $K \in \mathbb{Q}_0^+$. Then we define $Q' := Q + K$ as follows: $Q' = [q_*] \cup [(q'_{(a,b)})_{a,b \in \mathbb{N}}]$, where $q'_{(0,2)} := q_{(0,2)} + K$ and for all $(a, b) \neq (0, 2)$ $q'_{(a,b)} := q_{(a,b)}$. By definition, the annotation coefficient $q_{(0,2)}$ is the coefficient of the basic potential function $p_{(0,2)}(t) = \log(0|t| + 2) = 1$, so the annotation $Q + K$, adds cost K to the potential induced by Q . Further, we define the multiplication of an annotation Q by a constant K , denoted as $K \cdot Q$ pointwise. Moreover, let $P = [p_*] \cup [(p_{(a,b)})_{a,b \in \mathbb{N}}]$ be another annotation. Then the addition $P + Q$ of annotations P, Q is similarly defined pointwise.

5. Logarithmic Amortised Resource Analysis

In this section, we present the central contribution of this work. We delineate a novel type-and-effect system incorporating a potential-based amortised resource analysis capable of expressing *logarithmic* amortised costs. Soundness of the approach is established in Theorem 5.

Our potential-based amortised resource analysis is couched in a type system, given in Figure 5. If the type judgement $\Gamma | Q \vdash e : \alpha | Q'$ is derivable, then the worst-case cost of evaluating the expression e is bounded from above by the difference between the potential $\Phi(\sigma; \Gamma | Q)$ before the execution and the potential $\Phi(v | Q')$ of the value v obtained through the evaluation of the expression e . The typing system makes use of a *cost-free* semantics, which does not attribute any costs to the calculation. The cost-free typing judgement is denoted as $\Gamma | Q \vdash^{cf} e : \alpha | Q'$ and based on a cost-free variant of the application rule, denoted as (app : cf). The rule (app : cf) is defined as the rule (app), however, no costs are accounted for. Wrt. the cost-free semantics, the *empty signature*, denoted as $\alpha_1 \times \dots \times \alpha_n | \emptyset \rightarrow \beta | \emptyset$, is always admissible. We note that the cost-free signatures form a cone, as stated in the following remark:

Remark 10. If $\alpha | _ \rightarrow 1 | \times \dots \times \alpha_n | P \rightarrow \beta | P'$ and $\alpha | _ \rightarrow 1 | \times \dots \times \alpha_n | Q \beta Q'$ are both cost-free signatures for a function f , then any linear combination is admissible as cost-free signature of f , that is, we can assume $\alpha_1 \times \dots \times \alpha_n | K \cdot P + L \cdot Q \rightarrow \beta | K \cdot P' + L \cdot Q' \in F^{cf}(f)$, where $K, L \in \mathbb{Q}_0^+$.

$$\begin{array}{c}
 \frac{\forall c \geq 2 \ q_{(c)} = \sum_{a+b=c} q'_{(a,b)} \quad K = q'_*}{\emptyset | Q + K \vdash \mathbf{leaf} : \mathbb{T} | Q'} \text{ (leaf)} \qquad \frac{\Gamma | Q \vdash e : \alpha | Q' \quad K \geq 0}{\Gamma | Q + K \vdash e : \alpha | Q' + K} \text{ (shift)} \\
 \\
 \frac{\Gamma | R \vdash e : \beta | Q' \quad r_i = q_i \quad r_{(\vec{a},b)} = q_{(\vec{a},0,b)}}{\Gamma, x : \alpha | Q \vdash e : \beta | Q'} \text{ (w : var)} \qquad \frac{\Gamma, x : \alpha, y : \alpha | Q \vdash e[x, y] : \beta | Q'}{\Gamma, z : \alpha | \forall(Q) \vdash e[z, z] : \beta | Q'} \text{ (share)} \\
 \\
 \frac{q_1 = q_2 = q'_* \quad q_{(1,0,0)} = q_{(0,1,0)} = q'_* \quad q_{(a,a,c)} = q'_{(a,c)}}{x_1 : \mathbb{T}, x_2 : \mathbb{B}, x_3 : \mathbb{T} | Q \vdash (x_1, x_2, x_3) : \mathbb{T} | Q'} \text{ (node)} \\
 \\
 \frac{\circ \in \{<, >, =\}}{x_1 : \alpha, x_2 : \alpha | Q \vdash x_1 \circ x_2 : \mathbf{Bool} | Q} \text{ (cmp)} \qquad \frac{\Gamma | Q \vdash e_1 : \alpha | Q' \quad \Gamma | Q \vdash e_2 : \alpha | Q'}{\Gamma, x : \mathbf{Bool} | Q \vdash \mathbf{if } x \mathbf{ then } e_1 \mathbf{ else } e_2 : \alpha | Q'} \text{ (ite)} \\
 \\
 r_{(\vec{a},a,b)} = q_{(\vec{a},a,b)} \qquad r_{m+1} = r_{m+2} = q_{m+1} \\
 P_{(\vec{a},c)} = \sum_{a+b=c} q_{(\vec{a},a,b)} \qquad r_{(\vec{0},1,0,0)} = r_{(\vec{0},0,1,0)} = q_{m+1} \\
 \\
 \frac{\Gamma | P + q_{m+1} \vdash e_1 : \alpha | Q' \quad \Gamma, x_1 : \mathbb{T}, x_2 : \mathbb{B}, x_3 : \mathbb{T} | R \vdash e_2 : \alpha | Q' \quad q_i = r_i = p_i}{\Gamma, x : \mathbb{T} | Q \vdash \mathbf{match } x \mathbf{ with } | \mathbf{leaf} \rightarrow e_1 | \mathbf{leaf} \rightarrow e_2 : \alpha | Q'} \text{ (match)} \\
 \\
 p_i = q_i \quad p_{(\vec{a},c)} = q_{(\vec{a},\vec{0},c)} \quad r_j = q_{m+j} \quad r_{k+1} = p'_* \quad r_{(\vec{0},d,e)} = p'_{(d,e)} \\
 \forall \vec{b} \neq \vec{0} \left(r_{(\vec{b},0,0)} = q_{(\vec{0},\vec{b},0)} \right) \\
 \forall \vec{b} \neq \vec{0}, \vec{a} \neq \vec{0} \vee c \neq 0 \left(q_{(\vec{a},\vec{b},c)} = \sum_{(d,e)} p_{(\vec{a},c)}^{(\vec{b},d,e)} \right) \\
 \forall \vec{b} \neq \vec{0}, d \neq 0 \vee e \neq 0 \left(r_{(\vec{b},d,e)} = p_{(d,e)}^{(\vec{b},d,e)} \wedge \forall (d', e') \neq (d, e) \left(p_{(d',e')}^{(\vec{b},d,e)} = 0 \right) \wedge \right. \\
 \left. \wedge \sum_{(\vec{a},c)} p_{(\vec{a},c)}^{(\vec{b},d,e)} \geq p_{(d,e)}^{(\vec{b},d,e)} \wedge \forall \vec{a} \neq \vec{0} \vee c \neq 0 \left(p_{(\vec{a},c)}^{(\vec{b},d,e)} \neq 0 \rightarrow p_{(d,e)}^{(\vec{b},d,e)} \leq p_{(\vec{a},c)}^{(\vec{b},d,e)} \right) \right) \\
 \\
 \frac{\Gamma | P \vdash e_1 : \mathbb{T} | P' \quad \forall \vec{b} \neq \vec{0}, d \neq 0 \vee e \neq 0 \left(\Gamma | P^{(\vec{b},d,e)} \vdash^{\text{cf}} e_1 : \mathbb{T} | P^{(\vec{b},d,e)} \right) \quad \Delta, x : \mathbb{T} | R \vdash e_2 : \beta | Q'}{\Gamma, \Delta | Q \vdash \mathbf{let } x = e_1 \mathbf{ in } e_2 : \beta | Q'} \text{ (let : T)} \\
 \\
 p_i = q_i \quad p_{(\vec{a},c)} = q_{(\vec{a},\vec{0},c)} \qquad q_{(\vec{0},\vec{b},c)} = r_{(\vec{b},c)} \quad (\vec{b} \neq \vec{0}) \qquad r_j = q_{m+j} \\
 \Gamma | P \vdash e_1 : \alpha | \emptyset \qquad \Delta, x : \alpha | R \vdash e_2 : \beta | Q' \qquad \alpha \neq \mathbb{T} \\
 \Gamma, \Delta | Q \vdash \mathbf{let } x = e_1 \mathbf{ in } e_2 : \beta | Q' \text{ (let : gen)} \\
 \\
 \frac{\Gamma | P \vdash e : \alpha | P' \quad \Phi(\Gamma | P) \leq \Phi(\Gamma | Q) \quad \Phi(\Gamma | P') \geq \Phi(\Gamma | Q')}{\Gamma | Q \vdash e : \alpha | Q'} \text{ (w)} \qquad \frac{x \text{ a variable}}{x : \alpha | Q \vdash x : \alpha | Q} \text{ (var)} \\
 \\
 \frac{\alpha_1 \times \dots \times \alpha_n | P \rightarrow \beta | P' \in \mathcal{F}(f) \quad \alpha_1 \times \dots \times \alpha_n | Q \rightarrow \beta | Q' \in \mathcal{F}^{\text{cf}}(f) \quad K \in \mathbb{Q}^+}{x_1 : \alpha_1, \dots, x_n : \alpha_n | P + K \cdot Q \vdash f(x_1, \dots, x_n) : \beta | (P' + K \cdot Q') - 1} \text{ (app)}
 \end{array}$$

To ease notation, we set $\vec{a} := a_1, \dots, a_m, \vec{b} := b_1, \dots, b_k$ for vectors of indices $a_i, b_j \in \mathbb{N}$. Further, $i \in \{1, \dots, m\}, j \in \{1, \dots, k\}$ and $a, b, c, d, e \in \mathbb{N}$. Sequence elements of annotations, which are not constrained are set to zero.

Figure 5. Type system for logarithmic amortised resource analysis

Remark 11. Principally the type system can be parameterised in the resource metric (see, e.g., Hoffmann et al. 2012a). In this article, we focus on amortised and worst-case runtime complexity, symbolically measured through the number of function applications. It is straightforward to generalise this type system to other monotone cost models. Wrt. non-monotone costs, for example, heap usage, we expect the type system can also be readily adapted, but this is outside the scope of the article.

We consider the typing rules in turn; recall the convention that sequence elements of annotations are denoted by the lower-case letter of the annotation. Further, note that sequence elements which do not occur in any constraint are set to zero. The variable rule (var) types a variable of unspecified type α . As no actual costs are required, the annotation is unchanged. Similarly, no

resources are lost through the use of control operators. Hence, the definition of the rules (cmp) and (ite) is straightforward.

As exemplary constructor rules, we have rule (leaf) for the empty tree and rule (node) for the node constructor. Both rules define suitable constraints on the resource annotations to guarantee that the potential of the values is correctly represented.

The application rule (app) represents the application of a function given in P. Each application emits actual cost 1, which is indicated in the subtraction of 1. In its simplest form, that is, for the factor $K = 0$, the rule allows to directly read off the required annotation from the set of signatures \mathcal{F} . For arbitrary $K \in \mathbb{Q}_0^+$, the rule allows to combine some signature with cost with a cost-free signature. We note that Remark 10 would in fact allow us to add any positive linear combination of cost-free signatures; however, for performance reasons we refrain from doing so.

In the pattern matching rule (match) the potential freed through the destruction of the tree construction is added to the annotation R , which is used in the right premise of the rule. Note that $|R| = m + 2$, where m equals the number of tree types in the type context Γ .

The constraints expressed in the typing rules (let : T) and (let : gen) guarantee that the potential provided through annotation Q is distributed among the call to e_1 and e_2 , that is, this rule takes care of function composition. The numbers m, k , respectively, denote the number of tree types in Γ, Δ . Due to the sharing rule—discussed in a moment—we can assume wlog. that each variable in e_1 and e_2 occurs at most once.

First, consider the rule (let : gen), that is, the expression e_1 evaluates to a value w of arbitrary type $\alpha \neq \top$. In this case, the resulting value w cannot carry any potential. This is indicated through the empty annotation \emptyset in the typing judgement $\Gamma | P \vdash e_1 : \alpha | \emptyset$. Similarly, in the judgement $\Delta, x : \alpha | R \vdash e_2 : \beta | Q'$ for the expression e_2 , all available potential prior to the execution of e_2 stems from the potential embodied in the type context Δ wrt. annotation Q . This is enforced by the corresponding constraints. Suppose for $\vec{a} \neq \vec{0}$ and $\vec{b} \neq \vec{0}$, $q_{(\vec{a}, \vec{b}, c)}$ would be non-zero. Then the corresponding shared potential between the contexts Γ and Δ wrt. Q is discarded by the rule, as there is no possibility this potential is attached to the result type α .

Second, consider the more involved rule (let : T). To explain this rule, we momentarily assume that in Q no potential is shared, that is, $q_{(\vec{a}, \vec{b}, c)} = 0$, whenever $\vec{a} \neq \vec{0}, \vec{b} \neq \vec{0}$. In this sub-case, the rule can be simplified as follows:

$$\begin{array}{c}
 p_i = q_i \qquad \qquad \qquad r_j = q_{m+j} \\
 P_{(\vec{a}, c)} = q_{(\vec{a}, \vec{0}, c)} \quad r_{(\vec{b}, 0, c)} = q_{(\vec{0}, \vec{b}, c)} \quad (\vec{b} \neq \vec{0}) \quad r_{k+1} = p'_* \\
 \Gamma | P \vdash e_1 : \top | P' \quad \Delta, x : \top | R \vdash e_2 : \beta | Q' \quad r_{(\vec{0}, a, c)} = p'_{(a, c)} \\
 \hline
 \Gamma, \Delta | Q \vdash \mathbf{let} \ x = e_1 \ \mathbf{in} \ e_2 : \beta | Q' \quad (\mathbf{let} : T)
 \end{array}$$

Again the potential in Γ, Δ (wrt. annotation Q) is distributed for the typing of the expressions e_1, e_2 , respectively, which is governed by the constraints on the annotations. The simplified rule is obtained, as the assumption that no shared potential exists makes almost all constraints vacuous.

In particular, the cost-free derivation $\Gamma | P^{(\vec{b}, d, e)} \vdash^{cf} e_1 : \top | P'^{(\vec{b}, d, e)}$ is not required.

Finally, consider the most involved sub-case, where shared potentials are possible. Contrary to the simplified rules discussed above, such shared potential cannot be split between the type contexts Γ and Δ , respectively. Thus, the full rule necessarily employs the *cost-free semantics*.

Consequently, the premise $\Gamma | P^{(\vec{b}, d, e)} \vdash^{cf} e_1 : \alpha | P'^{(\vec{b}, d, e)}$ expresses that for all non-zero vectors \vec{b} and arbitrary indices d, e , the potentials $\Phi(\Gamma | P^{(\vec{b}, d, e)})$ suffice to cover the potential $\Phi(\alpha | P'^{(\vec{b}, d, e)})$, if no extra costs are emitted (compare Section 2). Intuitively, this represents that the values do not increase during the evaluation of e_1 to value w .

At last, the type system makes use of structural rules, like the *sharing* rule (share) and the weakening rules ($w : \mathbf{var}$) and (w). The *sharing* rule employs the sharing operator, defined in Lemma 8.

Note that the variables x, y introduced in the assumption of the typing rule are fresh variables that do not occur in Γ . Similarly, the rule (shift) allows to shift the potential before and after evaluation of the expression e by a constant K .

Note that the weakening rules embody changes in the potential of the type context of expressions considered. This amounts to the comparison on logarithmic expressions, principally a non-trivial task that cannot be directly represented as constraints in the type system. Instead, the rule (w) employs a symbolic potential expressions for these comparisons, replacing actual values for tree by variables. Let Γ denote a type context containing the type declarations $x_1 : T, \dots, x_m : T$ and let Q be an annotation of length m . Then the *symbolic potential*, denoted as $\Phi(\Gamma|Q)$, is defined as follows:

$$\Phi((x_1, \dots, x_m)|Q) := \sum_{i=1}^m q_i \cdot \text{rk}(x_i) + \sum_{(a_1, \dots, a_m), b \in \mathbb{N}} q_{((a_1, \dots, a_m), b)} \cdot P_{((a_1, \dots, a_m), b)}((x_1, \dots, x_m)),$$

where $P_{((a_1, \dots, a_m), b)}((x_1, \dots, x_m)) = \log(a_1 \cdot |x_1| + \dots + a_m \cdot |x_m| + b)$. To actually solve these constraints over symbolic potentials, we have to *linearise* the underlying comparisons of logarithmic expressions. This is taken up again in Section 7.

Definition 12. A program P is called *well-typed* if for any rule $f((x_1, \dots, x_k)) = e \in P$ and any annotated signature $\alpha_1 \times \dots \times \alpha_k | Q \rightarrow \beta | Q' \in \mathcal{F}(f)$, we have $x_1 : \alpha_1, \dots, x_k : \alpha_k | Q \vdash e : \beta | Q'$. A program P is called *cost-free well-typed*, if the cost-free typing relation is employed.

Before we state and prove the soundness of the presented type-and-effect system, we establish the following auxiliary result, employed in the correct assessment of the transfer of potential in the case of function composition, see Figure 5. See also the high-level description provided in Section 2.

Lemma 13. Assume $\sum_i q_i \log a_i \geq q \log b$ for some rational numbers $a_i, b > 0$ and $q_i \geq q$. Then, $\sum_i q_i \log(a_i + c) \geq q \log(b + c)$ for all $c \geq 1$.

Proof. Wlog, we can assume $q = 1$ and $q_i \geq 1$, as otherwise we simply divide the assumed inequality by q . Further, observe that the assumption $\sum_i q_i \log a_i \geq q \log b$ is equivalent to

$$\prod_i a_i^{q_i} \geq b. \tag{9}$$

First, we prove that

$$(x + y)^r \geq x^r + y^r \quad r \geq 1 \quad x, y \geq 0. \tag{10}$$

This is proved as follows. Fix some $x \geq 0$ and consider $(x + y)^r$ and $x^r + y^r$ as functions in y . It is then sufficient to observe that $(x + y)^r \geq x^r + y^r$ for $y = 0$ and that $\frac{d}{dy}(x + y)^r \geq \frac{d}{dy}(x^r + y^r)$ (the derivatives with regard to y) for all $y \geq 0$. Indeed, we have $\frac{d}{dy}(x + y)^r = r(x + y)^{r-1}$ and $\frac{d}{dy}(x^r + y^r) = ry^{r-1}$. Because of $r \geq 1$ and $x \geq 0$, we can thus deduce that $\frac{d}{dy}(x + y)^r \geq \frac{d}{dy}(x^r + y^r)$ for all $y \geq 0$.

Now we consider some $c \geq 1$. Combining (9) and (10), we get

$$\prod_i (a_i + c)^{q_i} \geq \prod_i (a_i^{q_i} + c^{q_i}) \geq \prod_i a_i^{q_i} + \prod_i c^{q_i} \geq b + c,$$

where we have used that $i \geq 1$, and that $q_i \geq 1$ and $c \geq 1$ imply $\prod_i c^{q_i} \geq c$. By taking the logarithm on both sides of the inequality, we obtain the claim. □

Finally, we obtain the following soundness result, which roughly states that if a program P terminates, then the difference in potential has paid its execution costs.²

Theorem (Soundness Theorem) *Let P be well-typed and let σ be a substitution. Suppose $\Gamma|Q \vdash e : \alpha|Q'$ and $\sigma \ell e \Rightarrow v$. Then $\Phi(\sigma; \Gamma|Q) - \Phi(v|Q') \geq \ell$. Further, if $\Gamma|Q \vdash^{cf} e : \alpha|Q'$, then $\Phi(\sigma; \Gamma|Q) \geq \Phi(v|Q')$.*

Proof. The proof embodies the high-level description given in Section 2. It proceeds by main induction on $\Pi : \sigma \ell e \Rightarrow v$ and by side induction on $\Xi : \Gamma|Q \vdash e : \alpha|Q'$, where the latter is employed in the context of the weakening rules. We consider only a few cases of interest. For example, for a case not covered: the variable rule (var) types a variable of unspecified type α . As no actual costs are required, the annotation is unchanged and the theorem follows trivially.

Case. Π derives $\sigma \stackrel{0}{\vdash} \mathbf{leaf} \Rightarrow \mathbf{leaf}$. Then Ξ consists of a single application of the rule (leaf):

$$\frac{\forall c \geq 2 \ q(c) = \sum_{a+b=c} q'(a,b)K = q'_*(\mathbf{leaf})}{\emptyset|Q + K \vdash \mathbf{leaf} : \Gamma|Q'}(\mathbf{leaf}).$$

By assumption, $Q = [(q(c))_{c \in \mathbb{N}}]$ is an annotation for the empty sequence of trees. On the other hand, $Q' = [(q'(a,b))_{a,b \in \mathbb{N}}]$ is an annotation of length 1. Note that $\text{rk}(\mathbf{leaf}) = 1$ by definition. Thus, we obtain

$$\begin{aligned} \Phi(\epsilon|Q + K) &= K + \sum_c q(c) \cdot \log(c) \\ &= K + \sum_{c \geq 2} q(c) \cdot \log(c) \\ &= q'_* + \sum_{a+b \geq 2} q'(a,b) \cdot \log(a+b) \\ &= q'_* + \sum_{a,b} q'(a,b) \cdot \log(a+b) \\ &= q'_* \text{rk}(\mathbf{leaf}) + \sum_{a,b} q'(a,b) p_{(a,b)}(\mathbf{leaf}) = \Phi(\mathbf{leaf}|Q'). \end{aligned}$$

Case. Suppose Π has the following from:

$$\frac{x_1 \sigma = t \quad x_2 \sigma = b \quad x_3 \sigma = u}{\sigma \stackrel{0}{\vdash} (x_1, x_2, x_3) \Rightarrow (t, b, u)}.$$

Wlog. Ξ consists of a single application of the rule (node):

$$\frac{q_1 = q_2 = q'_* \quad q_{(1,0,0)} = q_{(0,1,0)} = q'_* \quad q_{(a,a,b)} = q'(a,b)}{x_1 : \mathbb{T}, x_2 : \mathbb{B}, x_3 : \mathbb{T}|Q \vdash (x_1, x_2, x_3) : \mathbb{T}|Q'}(\mathbf{node})$$

By definition, we have $Q = [q_1, q_2] \cup [(q_{(a_1, a_2, b)})_{a_i, b \in \mathbb{N}}]$ and $Q' = [q'_*] \cup [(q'_{(a', b')})_{a', b' \in \mathbb{N}}]$. We set $\Gamma := x_1 : \mathbb{T}, x_2 : \mathbb{B}, x_3 : \mathbb{T}$ as well as $x_1 \sigma = u, x_2 \sigma = b$, and $x_3 \sigma = v$. Thus, $\Phi(\sigma; \Gamma|Q) = \Phi(u, v|Q)$ and

²A stated, soundness assumes termination of P , but our analysis is not restricted to terminating programs. To avoid the assumption the soundness theorem would have to be formulated wrt. to a partial big-step or a small step semantics, cf. Hoffmann and Hofmann (2010a); Moser and Schneckenreither (2020). We consider this outside the scope of this work.

we obtain

$$\begin{aligned} \Phi(u, v|Q) &= q_1 \cdot \text{rk}(u) + q_2 \cdot \text{rk}(v) + \sum_{a_1, a_2, b} q_{(a_1, a_2, b)} \cdot \log(a_1 \cdot |u| + a_2 \cdot |v| + b) \\ &\geq q'_* \cdot \text{rk}(u) + q'_* \cdot \text{rk}(v) + q_{(1,0,0)} \cdot \log(|u|) + q_{(0,1,0)} \cdot \log(|v|) + \\ &\quad + \sum_{a,b} q_{(a,a,b)} \cdot \log(a \cdot |u| + a \cdot |v| + b) \\ &= q'_* \cdot (\text{rk}(u) + \text{rk}(v) + \log(|u|) + \log(|v|)) + \\ &\quad + \sum_{a,b} q'_{(a,b)} \cdot \log(a \cdot (|u| + |v|) + b) \\ &= q'_* \cdot \text{rk}((u, b, v)) + \sum_{a,b} q'_{(a,b)} \cdot p_{(a,b)}((u, b, v)) = \Phi((u, b, v)|Q'). \end{aligned}$$

Case. Suppose $\sigma \ell e \Rightarrow v$ and let the last rule in Ξ be of the following form:

$$\frac{\Gamma|Q \vdash e : \alpha|Q'}{\Gamma|Q + K \vdash e : \alpha|Q' + K},$$

where $K \geq 0$. By SIH, we have that $\Phi(\sigma; \Gamma|Q) - \Phi(v|Q') \geq \ell$, from which we immediately obtain:

$$\Phi(\sigma; \Gamma|Q)K - \Phi(v|Q') - K = \Phi(\sigma; \Gamma|Q) - \Phi(v|Q') \geq \ell.$$

Case. Consider the first (match) rule, where Π ends as follows:

$$\frac{x\sigma = \text{leaf} \quad \sigma \upharpoonright^\ell e_1 \Rightarrow v}{\sigma \upharpoonright^\ell \text{match } x \text{ with } | \text{leaf} \rightarrow e_1 | (x_0, x_1, x_2) \rightarrow e_2 \Rightarrow v}.$$

Wlog. we may assume that Ξ ends with the related application of the (match) rule:

$$\frac{\begin{array}{l} r_{(\bar{a}, a, a, b)} = q_{(\bar{a}, a, b)} \quad r_{m+1} = r_{m+2} = q_{m+1} \\ P_{(\bar{a}, c)} = \sum_{a+b=c} q_{(\bar{a}, a, b)} \quad r_{(\bar{0}, 1, 0, 0)} = r_{(\bar{0}, 0, 1, 0)} = q_{m+1} \\ \Gamma|P + q_{m+1} \vdash e_1 : \alpha|Q' \quad \Gamma, x_1 : T, x_2 : B, x_3 : T|R \vdash e_2 : \alpha|Q' \quad q_i = r_i = p_i \end{array}}{\Gamma, x : T|Q \vdash \text{match } x \text{ with } | \text{leaf} \rightarrow e_1 | (x_1, x_2, x_3) \rightarrow e_2 : \alpha|Q'}.$$

Let Q be an annotation of length $m + 1$ while Q' is of length 1. Thus, annotations P, R have lengths $m, m + 2$, respectively. We write $\vec{t} := t_1, \dots, t_n$ for the substitution instances of the variables in Γ . Further $x\sigma = \text{leaf}$, where the latter equality follows from the assumption on Π . By definition and the constraints given in the rule, we obtain

$$\begin{aligned} \Phi(\sigma; \Gamma, x : T|Q) &= \sum_i q_i \text{rk}(t_i) + q_{m+1} \text{rk}(\text{leaf}) + \sum_{\bar{a}, a, c} q_{(\bar{a}, a, c)} \log(\bar{a}|\vec{t}| + a|\text{leaf}| + c) \\ &= \sum_i q_i \text{rk}(t_i) + q_{m+1} (\text{rk}(\text{leaf})) + \sum_{\bar{a}, a, c} q_{(\bar{a}, a, c)} \log(\bar{a}|\vec{t}| + a + c) \\ &= \Phi(\sigma; \Gamma|P)q_{m+1}. \end{aligned}$$

Thus, $\Phi(\sigma; \Gamma, x : T|Q) = \Phi(\sigma; \Gamma|P + q_{m+1})$ and the theorem follows by an application of MIH.

Now, consider the second (match) rule, that is, Π ends as follows:

$$\frac{x\sigma = (t, a, u)\sigma' \ell e_2 \Rightarrow v}{\sigma \ell \text{match } x \text{ with } | \text{leaf} \rightarrow e_1 | (x_0, x_1, x_2) \rightarrow e_2 \Rightarrow v}.$$

As above, we may assume that Ξ ends with the related application of the (match) rule. In this sub-case, the assumption on Π yields $t := x\sigma = (u, b, v)$. By definition and the constraints given in the rule, we obtain

$$\begin{aligned} \Phi(\sigma; \Gamma, x : T | Q) &= \sum_i q_i \text{rk}(t_i) + q_{m+1} \text{rk}((u, b, v)) + \sum_{\vec{a}, a, c} q_{(\vec{a}, a, c)} \log(\vec{a}|\vec{t}| + a(|u, b, v|) + c) \\ &= \sum_i q_i \text{rk}(t_i) + q_{m+1} (\text{rk}(u) + \log(|u|) + \log(|v|) + \text{rk}(v)) + \\ &\quad + \sum_{\vec{a}, a, c} q_{(\vec{a}, a, c)} \log(\vec{a}|\vec{t}| + a(|u| + |v|) + c) \\ &= \Phi(\sigma; \Gamma, x_1 : T, x_2 : B, x_3 : T | R), \end{aligned}$$

where we write $\vec{a}|\vec{t}|$ as shorthand to denote componentwise multiplication.

Thus, $\Phi(\sigma; \Gamma, x : T | Q) = \Phi(\sigma; \Gamma, x_1 : T, x_2 : B, x_3 : T | R)$ and the theorem follows by an application of MIH.

Case. Consider the (let) rule, that is, Π ends in the following rule:

$$\frac{\sigma \Big|_{\ell_1} e_1 \Rightarrow w \quad \sigma[x \mapsto w] \Big|_{\ell_2} e_2 \Rightarrow v}{\sigma \Big|_{\ell_1 + \ell_2} \text{let } x = e_1 \text{ in } e_2 \Rightarrow v},$$

where $\ell = \ell_1 + \ell_2$. First, we consider the sub-case, where the type of e_1 is an arbitrary type α but not of type T , that is, we assume that Ξ ends in the following application of the (let : gen)-rule

$$\frac{\begin{array}{l} p_i = q_i \quad p_{(\vec{a}, c)} = q_{(\vec{a}, \vec{0}, c)} \quad q_{(\vec{0}, \vec{b}, c)} = r_{(\vec{b}, c)} \quad (\vec{b} \neq \vec{0}) \quad r_j = q_{m+j} \\ \Gamma | P \vdash e_1 : \alpha | \emptyset \quad \Delta, x : \alpha | R \vdash e_2 : \beta | Q' \quad \alpha \neq T \end{array}}{\Gamma, \Delta | Q \vdash \text{let } x = e_1 \text{ in } e_2 : \beta | Q'} \quad [(\text{let : gen})].$$

Recall that $\vec{a} = a_1, \dots, a_n$, $\vec{b} = b_1, \dots, b_m$, $i \in \{1, \dots, m\}$, $j \in \{1, \dots, k\}$ and a_i, b_j, a, b, c, d, e are natural numbers. Further, the annotations Q, P, R are of length $m + k, m$ and k , respectively, while the corresponding resulting annotations Q', P' and R' are of length 1.

By definition and due to the constraints expressed in the typing rule, we have

$$\begin{aligned} \Phi(\sigma; \Gamma, \Delta | Q) &= \sum_i q_i \text{rk}(t_i) + \sum_j q_{m+j} \text{rk}(u_j) + \sum_{\vec{a}, \vec{b}, c} q_{(\vec{a}, \vec{b}, c)} \log(\vec{a}|\vec{t}| + \vec{b}|\vec{u}| + c) \\ \Phi(\sigma; \Gamma | P) &= \sum_i q_i \text{rk}(t_i) + \sum_{\vec{a}, c} q_{(\vec{a}, \vec{0}, c)} \log(\vec{a}|\vec{t}| + c) \\ \Phi(w | \emptyset) &= 0 \\ \Phi(\sigma; \Delta, x : \alpha | R) &= \sum_j q_{m+j} \text{rk}(u_j) + r_{k+1} \text{rk}(w) + \sum_{\vec{b}, a, c} q_{(\vec{0}, \vec{b}, c)} \log(\vec{b}|\vec{u}| + c), \end{aligned}$$

where we set $\vec{t} := t_1, \dots, t_m$ and $\vec{u} := u_1, \dots, u_k$, denoting the substitution instances of the variables in Γ, Δ , respectively. Thus, we obtain

$$\Phi(\sigma; \Gamma, \Delta | Q) \geq \Phi(\sigma; \Gamma | P) + \Phi(\sigma; \Delta, x : \alpha | R).$$

By main induction hypothesis, we conclude that $\Phi(\sigma; \Gamma | P) - \Phi(w | \emptyset) \geq \ell_1$ and $\Phi(\sigma; \Delta, x : \alpha | R) - \Phi(v | Q') \geq \ell_2$, from which the sub-case follows.

Second, we consider the more involved sub-case, where e_1 is of type T . Thus, wlog. Ξ ends in the following application of the (let : T)-rule.

$$\begin{array}{c}
 p_i = q_i \quad p_{(\vec{a},c)} = q_{(\vec{a},\vec{0},c)} \quad r_j = q_{m+j} \quad r_{k+1} = p'_* \quad r_{(\vec{0},d,e)} = p'_{(d,e)} \\
 \forall \vec{b} \neq \vec{0} \left(r_{(\vec{b},0,0)} = q_{(\vec{0},\vec{b},0)} \right) \\
 \forall \vec{b} \neq \vec{0}, \vec{a} \neq \vec{0} \vee c \neq 0 \left(q_{(\vec{a},\vec{b},c)} = \sum_{(d,e)} p_{(\vec{a},c)}^{(\vec{b},d,e)} \right) \\
 \forall \vec{b} \neq \vec{0}, d \neq 0 \vee e \neq 0 \left(r_{(\vec{b},d,e)} = p'_{(d,e)}^{(\vec{b},d,e)} \wedge \forall (d', e') \neq (d, e) \left(p'_{(d',e')}^{(\vec{b},d,e)} = 0 \right) \wedge \right. \\
 \left. \wedge \sum_{(\vec{a},c)} p_{(\vec{a},c)}^{(\vec{b},d,e)} \geq p'_{(d,e)}^{(\vec{b},d,e)} \wedge \forall \vec{a} \neq \vec{0} \vee c \neq 0 \left(p_{(\vec{a},c)}^{(\vec{b},d,e)} \neq 0 \rightarrow p'_{(d,e)}^{(\vec{b},d,e)} \leq p_{(\vec{a},c)}^{(\vec{b},d,e)} \right) \right) \\
 \Gamma | P \vdash e_1 : T | P' \forall \vec{b} \neq \vec{0}, d \neq 0 \vee e \neq 0 \left(\Gamma | P^{(\vec{b},d,e)} \vdash_{cf} e_1 : T | P'^{(\vec{b},d,e)} \right) \Delta, x : T | R \vdash e_2 : \beta | Q' \\
 \hline
 \Gamma, \Delta | Q \vdash \text{let } x = e_1 \text{ in } e_2 : \beta | Q' \quad (\text{let} : T),
 \end{array}$$

where the annotations Q, P, R, Q', P' and the sequences \vec{a}, \vec{b} are as above. Further, for each sequence $\vec{b} \neq \vec{0}$, $p^{(\vec{b},u,v)}$ and $p'^{(\vec{b},u,v)}$ denote annotations of length m . By definition and due to the constraints expressed in the typing rule, we have for all $\vec{b} \neq \vec{0}$:

$$\begin{aligned}
 \Phi(\sigma; \Gamma, \Delta | Q) &= \sum_i q_i \text{rk}(t_i) + \sum_j q_j \text{rk}(u_j) + \sum_{\vec{a} \neq \vec{0} \vee \vec{b} \neq \vec{0} \vee c \neq 0} q_{(\vec{a},\vec{b},c)} \log(\vec{a}|\vec{t}| + \vec{b}|\vec{u}| + c) \\
 \Phi(\sigma; \Gamma | P) &= \sum_i q_i \text{rk}(t_i) + \sum_{\vec{a} \neq \vec{0} \vee c \neq 0} q_{(\vec{a},\vec{0},c)} \log(\vec{a}|\vec{t}| + c) \\
 \Phi(w | P') &= r_{k+1} \text{rk}(w) + \sum_{a \neq 0 \vee c \neq 0} r_{(\vec{0},a,c)} \log(a|w| + c) \\
 \Phi(\sigma; \Gamma | P^{(\vec{b},d,e)}) &= \sum_{\vec{a} \neq \vec{0} \vee c \neq 0} p_{(\vec{a},c)}^{(\vec{b},d,e)} \log(\vec{a}|\vec{t}| + c) \\
 \Phi(w | P'^{(\vec{b},d,e)}) &= p'_{(d,e)}^{(\vec{b},d,e)} \log(d|w| + e) \\
 \Phi(\sigma; \Delta, x : T | R) &= \sum_j q_j \text{rk}(u_j) + r_{k+1} \text{rk}(w) + \sum_{\vec{a} \neq \vec{0} \vee d \neq 0 \vee e \neq 0} r_{(\vec{b},d,e)} \log(\vec{b}|\vec{u}| + d|w| + e),
 \end{aligned}$$

where we set $\vec{t} := t_1, \dots, t_m$ and $\vec{u} := u_1, \dots, u_k$, denoting the substitution instances of the variables in Γ, Δ , respectively.

By main induction hypothesis, we conclude that $\Phi(\sigma; \Gamma | P) - \Phi(w | P') \geq \ell_1$. Further, for all $\vec{b} \neq \vec{0}, d \neq 0 \vee e \neq 0$, we have, due to the cost-free typing constraints $\Phi(\sigma; \Gamma | P^{(\vec{b},d,e)}) \geq \Phi(w | P'^{(\vec{b},d,e)})$. The latter yields more succinctly (for all $\vec{b} \neq \vec{0}, d \neq 0 \vee e \neq 0$) that

$$\sum_{\vec{a},c} p_{(\vec{a},c)}^{(\vec{b},d,e)} \log(\vec{a}|\vec{t}| + c) \geq p'_{(d,e)}^{(\vec{b},d,e)} \log(d|w| + e). \tag{11}$$

A third application of MIH yields that $\Phi(\sigma; \Delta, x : T | R) - \Phi(v | Q') \geq \ell_2$. Due to the conditions $\sum_{(a,c)} p_{(\vec{a},c)}^{(\vec{b},d,e)} \geq p'_{(d,e)}^{(\vec{b},d,e)}$, for all $(d', e') \neq (d, e)$, $p'_{(d',e')}^{(\vec{b},d,e)} = 0$ and for all \vec{a}, c ($p_{(\vec{a},c)}^{(\vec{b},d,e)} \neq 0 \rightarrow p'_{(d,e)}^{(\vec{b},d,e)} \leq p_{(\vec{a},c)}^{(\vec{b},d,e)}$), we can apply Lemma 13 to Equation (11) and obtain

$$\sum_{\vec{a} \neq \vec{0} \vee c \neq 0} p_{(\vec{a},c)}^{(\vec{b},d,e)} \log(\vec{a}|\vec{t}| + \vec{b}|\vec{u}| + c) \geq p'_{(d,e)}^{(\vec{b},d,e)} \log(\vec{b}|\vec{u}| + d|w| + e).$$

Due to the condition $(q_{(\vec{a}, \vec{b}, c)} = \sum_{(d,e)} p_{(\vec{a}, c)}^{(\vec{b}, d, e)})$ for all $\vec{b} \neq \vec{0}, \vec{a} \neq \vec{0} \vee c \neq 0$, we can sum those equations for all $d \neq 0 \vee e \neq 0$ and obtain (for all $\vec{b} \neq \vec{0}, d \neq 0 \vee e \neq 0$) that

$$\sum_{\vec{a} \neq \vec{0} \vee c \neq 0} q_{(\vec{a}, \vec{b}, c)} \log(\vec{a}|\vec{t}| + \vec{b}|\vec{u}| + c) \geq \sum_{d \neq 0 \vee e \neq 0} r_{(\vec{b}, d, e)} \log(\vec{b}|\vec{u}| + d|w| + e).$$

We can combine the above fact to conclude the case.

Case. Finally, we consider the application rules (app) and (app : cf). As the cost-free variant only differs in so far that costs are not counted by (app : cf), it suffices to consider the rule (app). Let $f((x_1, \dots, x_k)) = e \in P$ and let Π derives $\sigma \stackrel{\ell+1}{\vdash} f((x_1, \dots, x_k)) \Rightarrow v$. We consider the costed typing $x_1 : \alpha_1, \dots, x_k : \alpha_k | (P + K \cdot Q) + 1 \vdash^{cf} f((x_1, \dots, x_k)) : \alpha | P' - 1 + K \cdot Q'$, where $K \in \mathbb{Q}_0^+$. Set $\Gamma := x_1 : \alpha_1, \dots, x_k : \alpha_k$. As P is well-typed, we have

$$\Gamma | P \vdash e : \beta | P' \quad \text{and} \quad \Gamma | Q \vdash^{cf} e : \beta | Q'.$$

We can apply MIH wrt. the evaluation Π' of $\sigma \stackrel{\ell}{\vdash} e \Rightarrow v$ to conclude $\Phi(\sigma; \Gamma | P) - \Phi(v | P') \geq \ell$ as well as $\Phi(\sigma; \Gamma | Q) \geq \Phi(v | Q')$. By monotonicity of addition and multiplication:

$$\begin{aligned} \Phi(\sigma; \Gamma | P + K \cdot Q) &= \Phi(\sigma; \Gamma | P) + K \cdot \Phi(\sigma; \Gamma | Q) \\ &\geq (\Phi(v | P') + \ell) + K \cdot \Phi(v | Q') = \Phi(v | P' + K \cdot Q') + \ell. \end{aligned}$$

Thus,

$$\begin{aligned} &\Phi(\sigma; \Gamma | P + K \cdot Q) - \Phi(v | P' - 1 + K \cdot Q') = \\ &= (\Phi(\sigma; \Gamma | P + K \cdot Q) - \Phi(v | P' + K \cdot Q')) + 1 \geq \ell + 1. \end{aligned}$$

From this, the case follows, which completes the proof of the soundness theorem. □

Remark 14. We note that the basic resource functions can be generalised to additionally represent linear functions in the size of the arguments. The above soundness theorem is not affected by this generalisation.

In the next section, we exemplify the use of the proposed type-and-effect system, cf. Figure 5, on the motivating example.

6. Analysis

As promised, we apply in this section the proposed type-and-effect system to obtain an optimal analysis of the amortised costs of the zig-zig case of *splaying*, once the type annotations are fixed. As a preparatory step and also to emphasise the need for the cost-free semantics, we make precise the informal account of compositional reasoning given in Section 2.

6.1 Let-normal form

We consider the expression $(al, a, (ar, b, (br, c, cr))) =: t'$, which becomes the following expression e in let-normal form:

```

1 let t''' = (br, c, cr) in (
2   let t'' = (ar, b, t''') in ((al, a, t''))
3 )
```

The expression e is typable with the following derivation, where the expression e' abbreviates $let\ t'' \ (ar, b, t''')\ in\ ((al, a, t''))$. (We have ignored expressions of base type to

increase readability.)

$$\frac{br: \mathbb{T}, cr: \mathbb{T} | Q_1 \vdash (br, c, cr): \mathbb{T} | Q'_1 \quad \frac{\frac{q_{(1,1,0)}^3 = q_{(1,0)}^3}{ar: \mathbb{T}, t''': \mathbb{T} | Q_3 \vdash (ar, b, t'''): \mathbb{T} | Q'_3}}{ar: \mathbb{T}, al: \mathbb{T}, t''': \mathbb{T} | Q_2 \vdash e': \mathbb{T} | Q'_2} (*)}{br: \mathbb{T}, cr: \mathbb{T}, ar: \mathbb{T}, al: \mathbb{T} | Q \vdash e: \mathbb{T} | Q'}$$

Here, we employ the derivability of the following type judgement (12) by a single application of (node), wrt. the annotation Q_4, Q' given below.

$$al: \mathbb{T}, t'': \mathbb{T} | Q_4 \vdash (al, a, t''): \mathbb{T} | Q'. \tag{12}$$

It is not difficult to check that the above derivation indeed proves well-typedness of the expression e wrt. the below given type annotations.

$$\begin{aligned} Q: q_1 = q_2 = q_3 = q_4 = 1; q_{(1,1,1,0,0)} = 1; q_{(1,1,0,0,0)} = 1; q_{(0,0,0,0,2)} = 1, \\ q_{(1,0,0,0,0)} = q_{(0,1,0,0,0)} = q_{(0,0,1,0,0)} = q_{(0,0,0,1,0)} = 1, \\ Q': q'_* = 1, q'_{(0,2)} = 1, \\ Q_1: q_1^1 = q_1 = 1; q_2^1 = q_2 = 1; q_{(0,0,2)}^1 = q_{(0,0,0,0,2)} = 1, \\ q_{(1,1,0)}^1 = q_{(1,1,0,0,0)}; q_{(1,0,0)}^1 = q_{(1,0,0,0,0)} = 1; q_{(0,1,0)}^1 = q_{(0,1,0,0,0)} = 1, \\ Q'_1: q'^1_* = 1; q'^1_{(1,0)} = 1; q'^1_{(0,2)} = 1, \\ P^{(1,0,1,0)}: p_{(1,1,0)}^{(1,0,1,0)} = q_{(1,1,1,0,0)} = 1, \\ P'^{(1,0,1,0)}: p'_{(1,0)}^{(1,0,1,0)} = 1, \\ Q_2: q_1^2 = q_3 = 1; q_2^2 = q_4 = 1; q_3^2 = q'^1_* = 1; , \\ q_{(1,0,0,0)}^2 = q_{(0,0,1,0,0)} = 1; q_{(0,1,0,0)}^2 = q_{(0,0,0,1,0)} = 1; , \\ q_{(0,0,1,0)}^2 = q'^1_{(1,0)} = 1; q_{(0,0,0,2)}^2 = q'^1_{(0,2)} = 1; q_{(1,0,1,0)}^2 = p'_{(1,0)}^{(1,0,1,0)} = 1, \\ Q_3: q_1^3 = q_1^2 = 1; q_2^3 = q_2^2 = 1; q_{(0,0,2)}^3 = q_{(0,0,0,2)}^2 = 1 \\ q_{(1,0,0)}^3 = q_{(1,0,0,0)}^2 = 1; q_{(0,1,0)}^3 = q_{(0,0,1,0)}^2 = 1; q_{(1,1,0)}^3 = q_{(1,0,1,0)}^2 = 1, \\ Q'_3: q'^3_* = 1, q'^3_{(1,0)} = 1, q'^3_{(0,2)} = 1 \\ Q_4: q_1^4 = q_2^2 = 1; q_{(0,0,2)}^4 = q'^3_{(0,2)} = 1; q_{(1,0,0)}^4 = q_{(0,1,0,0)}^2 = 1; q_{(0,1,0)}^4 = q_{(1,0)}^3 = 1 \end{aligned}$$

In the inference marked with (*), we employ the (almost trivial) correctness of the following cost-free typing derivation for $br: \mathbb{T}, cr: \mathbb{T} | P^{(1,0,1,0)} \vdash^{cf} (br, c, cr): \mathbb{T} | P'^{(1,0,1,0)}$. (For instantiation of the rule (let : \mathbb{T}) note $\vec{b} = (1, 0)$.)

$$\frac{p_{(1,1,0)}^{(1,0,1,0)} = p'_{(1,0)}^{(1,0,1,0)}}{br: \mathbb{T}, cr: \mathbb{T} | P^{(1,0,1,0)} \vdash^{cf} (br, c, cr): \mathbb{T} | P'^{(1,0,1,0)}}. \tag{13}$$

For all $\vec{b} \neq (0)$, $\vec{b} \neq (1)$ and arbitrary d, e , we set $P(\vec{b}, d, e) = P'(\vec{b}, d, e) := \emptyset$. Our prototype fully automatically checks correctness of the above given annotations.

We emphasise that the involved (let)-rule, employed in step (*) cannot be avoided. In particular, the additional cost-free derivation (13) is essential. Observe the annotation marked in red in the calculation above. Eventually, these amount to a shared potential employed in step (*). The cost-free semantics allows us to exploit this shared potential, which otherwise would have to be discarded.

To wit, assume momentarily the rule (let) would not make use of cost-free reasoning, similar to the simplified (let)-rule, that we have used in the explanations on page 14. Then the shared potential represented by the coefficient $q_{(1,1,1,0,0)} \in Q$ is discarded by the rule. However, this potential is then missing, if we attempt to type the judgement

$$ar : T, al : T, t''' : T \mid R \vdash e' : T \mid Q'$$

where R is defined as Q_2 , except that $r_{(1,0,1,0)} = 0$. Thus, this attempt fails. (Note that the corresponding coefficient of Q_2 , marked in red, is non-zero.)

Remark 15. To some extent, this is in contrast to the use of cost-free semantics in the literature (Hoffmann, 2011; Hoffmann *et al.*, 2012a; Hofmann and Moser 2015; Hoffmann *et al.*, 2017; Moser and Schneckenreither, 2020). While cost-free semantics appear as an add-on in these works, essential only if we want to capture non-tail-recursive programs, cost-free semantics is essential in our context—it is already required for the representation of simple values.

6.2 Splay trees

In this subsection, we exemplify the use of the type system presented in the last section on the function `splay`, cf. Figure 2. Our amortised analysis of splaying yields that the amortised cost of `splay a t` is bound by $3 \log(|t|) + 1$, where the actual cost counts the number of recursive calls to `splay`, cf. Sleator and Tarjan (1985); Schoenmakers (1993); Nipkow (2015). To verify this amortised cost, we derive

$$a : \text{Base}, t : T \mid Q \vdash e : T \mid Q' \tag{14}$$

where the expression e is the definition of `splay` given in Figure 2 and the annotations Q and Q' are as follows:

$$\begin{aligned} Q : q_1 = 1, q_{(1,0)} = 3, q_{(0,2)} = 1, \\ Q' : q'_* = 1. \end{aligned}$$

Remark that the amortised cost of splaying is represented by the coefficients $q_{(1,0)}$ and $q_{(0,2)}$, expressing in sum $3 \log(|t|) + 1$. Note, further that the coefficient q_1, q'_* , represent Schoenmakers' potential, that is, $\text{rk}(t)$ and $\text{rk}(\text{splay } a \ t)$, respectively.

We restrict to the zig-zig case: $t = ((bl, b, br, c, cr))$ together with the recursive call `splay a bl = (al, a', ar)` and $a < b < c$. Thus, `splay a t` yields $(al, a', (ar, b, (br, c, cr))) =: t'$. Recall that a need not occur in t ; in this case, the last element a' before a leaf was found is rotated to the root. Our prototype checks correctness of these annotations automatically.

Let e_1 denote the subexpression of the definition of splaying, starting in program line 4. On the other hand, let e_2 denote the subexpression defined from lines 5–15 and let e_3 denote the program code within e_2 starting in line 8. Finally, the expression in lines 11 and 12 expand to the following, if we remove part of the syntactic sugar:

$$e_4 := \text{let } x = \text{splay } a \ bl \ \text{in match } x \ \text{with } | \text{leaf} \rightarrow \text{leaf} \mid (al, a', ar) \rightarrow t'.$$

Below, we show a simplified derivation of (14), where we have focused only on a particular path in the derivation tree, suited to the considered zig-zig case of the definition of splaying. Omission of premises or rules is indicated by double lines in the inference step. Again we make crucial use of the cost-free semantics in this derivation. The corresponding inference step is marked with (*) and the employed shared potentials are marked in red.

$$\frac{\frac{\frac{\text{splay: } \Gamma|Q \rightarrow \Gamma|Q'}{a: B, bl: \Gamma|Q \vdash \text{splay } a \text{ bl: } \Gamma|Q' - 1} \quad \frac{\frac{\frac{\frac{\Delta, cr: \Gamma, br: \Gamma, al: \Gamma, a': B, ar: \Gamma|Q_5 \vdash t': \Gamma|Q'}{\Delta, cr: \Gamma, br: \Gamma, x: \Gamma|Q_4 \vdash \text{match } x \text{ with } |(al, a', ar) \rightarrow t': \Gamma|Q'}{\Gamma, cr: \Gamma, bl: \Gamma, br: \Gamma|Q_3 \vdash e_4: \Gamma|Q'}{\Gamma, cr: \Gamma, bl: \Gamma, br: \Gamma|Q_3 \vdash e_3: \Gamma|Q'} \text{ (w)}}{\Gamma, cr: \Gamma, bl: \Gamma, br: \Gamma|Q_2 \vdash e_3: \Gamma|Q'}}{a: B, b: B, cl: \Gamma, cr: \Gamma|Q_1 \vdash \text{match } cl \text{ with } |leaf \rightarrow (cl, c, cr)| (bl, b, br) \rightarrow e_3: \Gamma|Q'}}{a: B, cl: \Gamma, c: B, cr: \Gamma|Q_1 \vdash \text{if } a = c \text{ then } (cl, c, cr) \text{ else } e_2: \Gamma|Q'}}{a: B, t: \Gamma|Q \vdash \text{match } t \text{ with } |leaf \rightarrow leaf| (cl, c, cr) \rightarrow e_1: \Gamma|Q'} \text{ (*)}$$

We abbreviate $\Gamma := a: B, b: B, c: B$, $\Delta := b: B, c: B$. In addition to the original signature of splaying, $B \times \Gamma|Q \rightarrow \Gamma|Q'$, we use the following annotations, induced by constraints in the type system, cf. Figure 5. As in Section 6.1, we mark annotations that require cost-free derivations in the (let : Γ) rule in red.

$$\begin{aligned}
 Q_1 : q_1^1 &= q_2^1 = q_1 = 1, q_{(1,1,0)}^1 = q_{(1,0)} = 3, q_{(1,0,0)}^1 = q_{(0,1,0)}^1 = q_1 = 1, q_{(0,0,2)}^1 = q_{(0,2)} = 1, \\
 Q_2 : q_1^2 &= q_2^2 = q_3^2 = 1, q_{(0,0,0,2)}^2 = 1, q_{(1,1,1,0)}^2 = q_{(1,1,0)}^1 = 3, q_{(0,1,1,0)}^2 = q_{(1,0,0)}^1 = 1, \\
 & q_{(1,0,0,0)}^2 = q_{(0,1,0)}^1 = 1, q_{(0,1,0,0)}^2 = q_{(0,0,1,0)}^2 = q_1^1 = 1, \\
 Q_3 : q_1^3 &= q_2^3 = q_3^3 = 1, q_{(0,0,0,2)}^3 = 2, \\
 & q_{(0,1,0,0)}^3 = 3, q_{(1,0,0,0)}^3 = q_{(0,0,1,0)}^3 = q_{(1,0,1,0)}^3 = q_{(1,1,1,0)}^3 = 1.
 \end{aligned}$$

In the step marked with the rule (w) in the derivation above, a *weakening* step is applied, which amounts to the following inequality:

$$\Phi(\Gamma, cr: \Gamma, bl: \Gamma, br: \Gamma|Q_2) \geq \Phi(\Gamma, cr: \Gamma, bl: \Gamma, br: \Gamma|Q_3).$$

We emphasise that this step can neither be avoided nor easily moved to the axioms of the derivation. We verify the correctness of *weakening* through a direct comparison. Let σ be a substitution. Then, we have

$$\begin{aligned}
 \Phi(\sigma; cr: \Gamma, bl: \Gamma, br: \Gamma|Q_2) &= 1 + \text{rk}(cr) + \text{rk}(bl) + \text{rk}(br) + 3 \log(|cr| + |bl| + |br|) + \\
 & \quad + \log(|bl| + |br|) + \log(|cr|) + \log(|bl|) + \log(|br|) \\
 &= 1 + \text{rk}(cr) + \text{rk}(bl) + \text{rk}(br) + 2 \log(|t|) + \log(|t|) + \\
 & \quad + \log(|bl| + |br|) + \log(|cr|) + \log(|bl|) + \log(|br|) \\
 &\geq 1 + \text{rk}(cr) + \text{rk}(bl) + \text{rk}(br) + \log(|bl|) + \log(|br| + |cr|) + 2 + \\
 & \quad + \log(|bl| + |br| + |cr|) + \\
 & \quad + \log(|bl| + |br|) + \log(|cr|) + \log(|bl|) + \log(|br|) + \\
 &\geq \text{rk}(bl) + 1 + 3 \log(|bl|) + \text{rk}(cr) + \text{rk}(br) + \log(|br|) + \\
 & \quad + \log(|cr|) + \log(|br| + |cr|) + \\
 & \quad + \log(|bl| + |br| + |cr|) + 1 = \Phi(\sigma; cr: \Gamma, bl: \Gamma, br: \Gamma|Q_3).
 \end{aligned}$$

Note that we have used Lemma 16 in the third line to conclude

$$2 \log(|t|) \geq \log(|bl|) + \log(|br| + |cr|) + 2,$$

as we have $|t| = |((bl, b, br), c, cr)| = |bl| + |br| + |cr|$. Furthermore, we have only used monotonicity of log and formal simplifications.

Further, we verify the use of the (let : T)-rule, marked with (*) in the proof. Consider the following annotation Q_4 :

$$\begin{aligned}
 Q_4: q_1^4 &= q_1^3; q_2^4 = q_3^3; q_3^4 = q' * : q_{(1,0,0,0)}^4 = q_{(1,0,0,0)}^3; q_{(0,1,0,0)}^4 = q_{(0,0,1,0)}^3; \\
 q_{(1,1,0,0)}^4 &= q_{(1,0,1,0)}^3; q_{(1,1,1,0)}^4 = p'_{(1,0)}^{(1,1,1,0)} = 1 \\
 P^{(1,1,1,0)}: p_{(1,0)}^{(1,1,1,0)} &= q_{(1,1,1,0)}^3 = 1 \\
 P'^{(1,1,1,0)}: p'_{(1,0)}^{(1,1,1,0)} &= 1
 \end{aligned}$$

To see that Q_4 is consistent with the constraints on resource annotations in the (let : T)-rule, we first note that

$$Q + 1: q = q_2^3 = 1, q_{(1,0)} = q_{(0,1,0,0)}^3 = 3; q_{(0,2)} = q_{(0,0,0,2)}^3.$$

Hence, the constraints on the annotations for the left typing tree in the (let : T)-rule amount to the following:

$$q = q_2^3 = 1 \quad q_{(1,0)} = q_{(0,1,0,0)}^3 = 3 \quad q_{(0,2)} = q_{(0,0,0,2)} = 2 \quad q' * = q_3^4 = 1,$$

which are fulfilled. Further, the right typing tree yields the constraints:

$$\begin{aligned}
 q_1^4 = q_1^3 = 1 \quad q_2^4 = q_3^3 = 1 \quad q_{(1,0,0,0)}^4 = q_{(1,0,0,0)}^3 = 1 \quad q_{(0,1,0,0)}^4 = q_{(0,0,1,0)}^3 = 1 \\
 q_{(1,1,0,0)}^4 = q_{(1,0,1,0)}^3 = 1,
 \end{aligned}$$

which are also fulfilled. Hence, it remains to check the correctness of the constraints for the actual cost-free derivation. First, note that for the vector $\vec{b} = (1, 1)$, the cost-free derivation needs to be checked wrt. the annotation pair $P^{(1,1,1,0)} = [p_{(1,0)}^{(1,1,1,0)}]$ and $P'^{(1,1,1,0)} = [p'_{(1,0)}^{(1,1,1,0)}]$. Second, the various constraints in the rule (let : T) simplify to the inequality $p_{(1,0)}^{(1,1,1,0)} \geq p'_{(1,0)}^{(1,1,1,0)}$, which holds. Third, the actual cost-free typing derivation reads as follows:

$$a : B, bl : T | P^{(1,1,1,0)} \vdash \text{splay } a \text{ } bl : T | P'^{(1,1,1,0)}. \tag{15}$$

The typing judgement (15) is derivable if the following cost-free signatures are employed for *splaying*:

$$\text{splay} : T | P \rightarrow^{cf} T | P' \quad T | \emptyset \rightarrow^{cf} T | \emptyset,$$

where $P = [p_{(1,0)}]$, $P' = [p'_{(1,0)}]$, with $p_{(1,0)} = p'_{(1,0)} := 1$. Recall that \emptyset denotes the empty annotation, where all coefficients are set to zero. By definition, $P = P^{(1,1,1,0)}$ and $P' = P'^{(1,1,1,0)}$. Informally, this cost-free signature is admissible, as the following equality holds:

$$\Phi(\sigma; a : B, bl : T | P) = \log(|bl|) = \log(|(al, a', ar)|) = \Phi((al, a', ar) : T | P').$$

Recall that we have $\text{splay } a \text{ } bl = (al, a', ar)$ for the recursive call and that $|bl| = |(al, a', ar)|$. As depicted below, the type derivation of (15) proceeds similarly to the derivation above in conjunction with the analysis in Subsection 6.1.

$$\begin{array}{c}
 \frac{\text{splay} : T | \emptyset \rightarrow T | \emptyset}{a : B, bl : T | \emptyset \vdash \text{splay } a \text{ } bl : T | \emptyset} \quad \frac{\Delta, cr : T, br : T, al : T, a' : B, ar : T | P_4 \vdash t' : T | P'}{\Delta, cr : T, br : T, x : T | P_3 \vdash \text{match } x \text{ with } |(al, a', ar)| \rightarrow t' : T | P'} \\
 \frac{\Gamma, cr : T, bl : T, br : T | P_2 \vdash e_4 : T | P'}{\Gamma, cr : T, bl : T, br : T | P_2 \vdash e_3 : T | P'} \quad (*) \\
 \frac{a : B, b : B, cl : T, cr : T | P_1 \vdash \text{match } cl \text{ with } |leaf| \rightarrow (cl, c, cr) | (bl, b, br) \rightarrow e_3 : T | P'}{a : B, cl : T, c : B, cr : T | P_1 \vdash \text{if } a = c \text{ then } (cl, c, cr) \text{ else } e_2 : T | P'} \\
 \frac{a : B, t : T | P \vdash \text{match } t \text{ with } |leaf| \rightarrow leaf | (cl, c, cr) \rightarrow e_1 : T | P'}{a : B, t : T | P \vdash \text{match } t \text{ with } |leaf| \rightarrow leaf | (cl, c, cr) \rightarrow e_1 : T | P'}
 \end{array}$$

As indicated, the cost-free derivation also requires the use of the full version of the rule (let : T), as marked by (*). In particular, the informal argument on the size of the argument and the result of splaying is built into the type system. We use the following annotations:

$$\begin{aligned}
 P &: p_{(1,0)} = 1 & P' &: p'_{(1,0)} = 1 \\
 P_1 &: P^1_{(1,1,0)} = p_{(1,0)} = 1 \\
 P_2 &: P^2_{(1,1,1,0)} = P^1_{(1,1,0)} = 1 \\
 P_3 &: P^3_{(1,1,1,0)} = p'_{(1,0)} = 1 \\
 P_4 &: P^4_{(1,1,1,1,0)} = P^3_{(1,1,1,0)} = 1
 \end{aligned}$$

Finally, one further application of the (match)-rule yields the desired derivation for suitable Q₅. See also the previous subsection. Note that one further application of the *weakening* rule is required here.

7. Linearisation and Expert Knowledge

In the context of the presented type-and-effect system (see Figure 5), an obvious challenge is the requirement to compare potentials symbolically (see Section 5) rather than compare annotations directly. This is in contrast to results on resource analysis for constant amortised costs, see, for example, Jost et al. (2009, 2010); Hoffmann et al. (2012a, 2017); Jost et al. (2017). Furthermore, the presence of logarithmic basic functions seems to necessitate the embodiment of nonlinear arithmetic. In particular, we need to make use of basic laws of the log functions, as the following one. A variant of the below fact has already been observed by Okasaki, cf. Okasaki (1999).

Lemma 16. *Let $x, y \geq 1$. Then $2 + \log(x) + \log(y) \leq 2 \log(x + y)$.*

Proof. We observe

$$(x + y)^2 - 4xy = (x - y)^2 \geq 0.$$

Hence $(x + y)^2 \geq 4xy$ and from the monotonicity of log we conclude $\log(xy) \leq \log(\frac{(x+y)^2}{4})$. By elementary laws of log, we obtain:

$$\log\left(\frac{(x + y)^2}{4}\right) = \log\left(\left(\frac{x + y}{2}\right)^2\right) = 2 \log(x + y) - 2,$$

from which the lemma follows as $\log(xy) = \log(x) + \log(y)$. □

A refined and efficient approach which targets linear constraints is achievable as follows. All logarithmic terms, that is, terms of the form $\log(\cdot)$ are replaced by new variables, focusing on finitely many. For the latter, we exploit the condition that in resource annotations only finitely many coefficients are non-zero. Consider the following inequality as prototypical example. Validity of the constraint ought to incorporate the monotonicity of log.

$$a_1 \log(|t|) + a_2 \log(|cr|) \geq b_1 \log(|t|) + b_2 \log(|cr|), \tag{16}$$

where we assume $t = \langle cl, c, cr \rangle$ for some value c and thus $|t| \geq |cr|$, cf. Section 6.2. Replacing $\log(|t|)$, $\log(|cr|)$ with new unknowns x, y , respectively, we represent (16) as follows:

$$\forall x, y \geq 0. a_1x + a_2y \geq b_1x + b_2y, \tag{17}$$

Here we keep the side-condition $x \geq y$ and observe that the unknowns x, y can be assumed to be non-negative, as they represent values of the log function. Thus, properties, for example, monotonicity of log, as well as properties like Lemma 16 above, can be expressed as inequalities over the introduced unknowns. For example, the inequality $x \geq y$ above represents the axiom of

monotonicity $\log(|t|) \geq \log(|cr|)$. All such obtained inequalities are collected as *expert or prior knowledge*. This entails that (17) is equivalent to the following existential constraint satisfaction problem:

$$\exists c, d. a_1 \geq b_1 + c \wedge a_2 \geq d \wedge b_2 \leq c + d. \tag{18}$$

We seek to systematise the derivation of inequalities such as (18) from expert knowledge. For that, we assume that the gathered prior knowledge is represented by a system of inequalities as $A\vec{x} \leq \vec{b}, \vec{x} \geq 0$, where A denotes a matrix with as many rows as we have prior knowledge, \vec{b} a column vector and \vec{x} the column vector of unknowns of suitable length; $\vec{x} \geq 0$ because log evaluates to non-negative values.

Below we discuss a general method for the derivation of inequalities such as (18) based on the affine form of Farkas’ Lemma. First, we state the variant of Farkas’ Lemma that we use in this article, cf. Schrijver (1999). Note that \vec{u} and \vec{f} denote column vectors of suitable length.

Lemma 17. (Farkas’ Lemma) *Suppose $A\vec{x} \leq \vec{b}, \vec{x} \geq 0$ is solvable. Then for all vectors \vec{u} and scalars λ , the following assertions are equivalent.*

$$\forall \vec{x} \geq 0. A\vec{x} \leq \vec{b} \Rightarrow \vec{u}^T \vec{x} \leq \lambda \tag{19}$$

$$\exists \vec{f} \geq 0. \vec{u}^T \leq \vec{f}^T A \wedge \vec{f}^T \vec{b} \leq \lambda \tag{20}$$

Proof. It is easy to see that from (20), we obtain (19). Assume (20). Assume further that $A\vec{x} \leq \vec{b}$ for some column vector \vec{x} . Then we have

$$\vec{u}^T \vec{x} \leq \vec{f}^T A \vec{x} \leq \vec{f}^T \vec{b} \leq \lambda.$$

Note that for this direction the assumption that $A\vec{x} \leq \vec{b}, \vec{x} \geq 0$ is solvable is not required.

With respect to the opposite direction, we assume (19). By assumption, $A\vec{x} \leq \vec{b}, \vec{x} \geq 0$ is solvable. Hence, maximisation of $\vec{u}^T \vec{x}$ under the side condition $A\vec{x} \leq \vec{b}, \vec{x} \geq 0$ is feasible. Let w denote the maximal value. Due to (19), we have $w \leq \lambda$.

Now, consider the dual asymmetric linear program to minimise $\vec{y}^T \vec{b}$ under side condition $\vec{y}^T A = \vec{u}^T$ and $\vec{y} \geq 0$. Due to the Dualisation Theorem, the dual problem is also solvable with the same solution

$$\vec{y}^T \vec{b} = \vec{u}^T \vec{x} = w.$$

We define $\vec{f} := \vec{y}$, which attains the optimal value w , such that $\vec{f}^T A = \vec{u}^T$ and $\vec{f} \geq 0$ such that $\vec{f}^T \vec{b} = w \leq \lambda$. This yields (20). □

Second, we discuss a method for the derivation of inequalities such as (18) based on Farkas’ Lemma. Our goal is to automatically discharge symbolic constraints such as $\Phi(\Gamma|P) \leq \Phi(\Gamma|Q)$ —as well as $\Phi(\Gamma|P') \geq \Phi(\Gamma|Q')$ —as required by the *weakening* rule (w) (see Section 5).

According to the above discussion, we can represent the inequality $\Phi(\Gamma|P) \leq \Phi(\Gamma|Q)$ by

$$\vec{p}^T \vec{x} + c_p \leq \vec{q}^T \vec{x} + c_q,$$

where \vec{x} is a finite vector of variables representing the base potential functions, \vec{p} and \vec{q} are column vectors representing the unknown coefficients of the non-constant potential functions, and c_p and c_q are the coefficients of the constant potential functions. We assume the expert knowledge is given by the constraints $A\vec{x} \leq \vec{b}, \vec{x} \geq 0$. We now want to derive conditions for \vec{p}, \vec{q}, c_p and c_q such that we can guarantee

$$\forall \vec{x} \geq 0. A\vec{x} \leq \vec{b} \Rightarrow \vec{p}^T \vec{x} + c_p \leq \vec{q}^T \vec{x} + c_q. \tag{21}$$

By Farkas’ Lemma, it is sufficient to find coefficients $\vec{f} \geq 0$ such that

$$\vec{p}^T \leq \vec{f}^T A + \vec{q}^T \wedge \vec{f}^T \vec{b} + c_p \leq c_q. \tag{22}$$

Hence, we can ensure Equation (21) by Equation (22) using the new unknowns \vec{f} .

We illustrate Equation (22) on the above example. We have $A = (-1\ 1)$, $b = 0$, $\vec{p} = (b_1\ b_2)^T$, $\vec{q} = (a_1\ a_2)^T$ as well as $c_p = c_q = 0$. Then, the inequality $fb + c_p \leq c_q$ simplifies to $0 \leq 0$ and can in the following be omitted. With the new unknown $f \geq 0$, we have

$$(b_1\ b_2) \leq f(-1\ 1) + (a_1\ a_2),$$

which we can rewrite to

$$b_1 \leq -f + a_1 \wedge b_2 \leq f + a_2,$$

easily seen to be equivalent to Equation (18).

Thus, the validity of constraints incorporating the monotonicity of log becomes expressible in a systematic way. Further, the symbolic constraints enforced by the *weakening* rule can be discharged systematically and become expressible as existential constraint satisfaction problems. Note that the incorporation of Farkas' Lemma in the above form subsumes the well-known practice of coefficient comparison for the synthesis of polynomial interpretations (Contejean et al., 2005), ranking functions (Podelski and Rybalchenko, 2004) or resource annotations in the context of constant amortised costs (Hoffmann et al., 2012a).

In the next section, we briefly detail our implementation of the established logarithmic amortised resource analysis, based on the observations in this section.

8. Implementation

Based on the principal approach, delineated in Section 2, we have provided a prototype implementation of the logarithmic amortised resource analysis detailed above. The prototype is capable of *type checking* a given resource annotation and requires user interaction to specify the structural inferences sharing and weakening. These can be applied manually to improve efficiency of type checking. In future work, we will strive for full automation, capable of *type inference*. In this section, we briefly indicate the corresponding design choices and heuristics used. Further, we present restrictions and future development areas of the prototype developed.

Template Potential Functions. Our potential-based method employs linear combinations of basic potential functions \mathcal{B} , cf. Definition 2. To fix the cardinality of the set of resource functions to be considered, we restrict the coefficients of the potential functions $p_{((a_1, \dots, a_m), b)}$. For the non-constant part, we demand that $a_i \in \{0, 1\}$, while the coefficients b , representing the constant part are restricted to $\{0, 1, 2\}$. This restriction to a relative small set of basic potential functions suitably controls the number of constraints generated for each inference rule in the type-and-effect system.

Type-and-Effect System. Following ideas of classical Hindley–Milner type inference, we collect for each node in the abstract syntax tree (AST) of the given program the constraints given by the corresponding inference rules in the type system (see Figure 5). As a pre-requisite, we restrict ourselves to three type annotations employed for each function symbol. (i) One indeterminate type annotation representing a function call with costs; (ii) one indeterminate cost-free type annotation to represent a zero-cost call; and one fixed cost-free annotation with the empty annotation that does not carry any potential. (iii) These restrictions were sufficient to handle the zig-zig case of splaying. A larger, potentially infinite set of type annotations is conceivable, as long as well-typedness is respected, cf. Definition 12. As noted in the context of the analysis of constant amortised complexity, an enlarged set of type annotations may be even required to handle non-tail recursive programs, cf. Hoffmann et al. (2012a, 2017). The collected constraints on the type annotations are passed to an SMT solver, in our case the SMT solver z3³ and solved over the positive rational numbers. Here we can directly encode the equalities and inequalities of the

³See <https://github.com/Z3Prover/z3/wiki>.

constraints given in the type system. Due to the use of Farkas' Lemma (Lemma 17) only linear constraints are generated.

Our implementation currently only supports type checking, taking user guidance into account and thus is semi-automated. While deriving constraints for the AST nodes of the program is straightforward (as there is only one type rule for every syntactic statement of our programming language), we currently require user interaction for the application of the structural rules.

Structural Rules. The structural rules can in principle be applied at every AST node of the program under analysis. However, they introduce additional variables and constraints and for performance reasons it is better to apply them sparingly. For the *sharing* rule we proceed as follows: We recall that the sharing rule allows us to assume that the type system is linear. In particular, we can assume that every variable occurs exactly once in the type context, which is exploited in the definition of the *let* rules. However, such an eager application of the sharing rule would directly yield to a size explosion in the number of constraints, as the generation of each fresh variable requires the generation of exponentially many annotations. Hence, we only apply sharing only when strictly necessary. In this way the typing context can be kept small. Similar to the sharing rule (*share*), variable weakening ($w : \text{var}$) is employed only when required. This in turn reduces the number of constraints generated.

For the *weakening* rule, we employ our novel methods for symbolically comparing logarithmic expressions, which we discussed in Section 7. Because of our use of Farkas' Lemma, weakening introduces new unknown coefficients, which again may result in a forbiddingly large search space. Thus, weakening steps are particularly costly. For performance reasons, we need to control the size of the resulting constraint system. Currently, we rely on the user to specify the number and place of the applications of the weakening rule. This is achieved through the provision of suitable *tactics* for type checking. Note that the weakening rule may need to be applied in the middle of a type derivation, see for example the typing derivation for our motivating example on page 23. This contrasts to the literature where the weakening rule can typically be incorporated into the axioms of the type system and thus dispensed with. Perhaps, a similar approach is possible in the context of logarithmic amortised resource analysis, but our current understanding does not support this.

Expert Knowledge. In Section 7, we propose the generation of a suitable matrix A collecting the *expert or prior knowledge* on such inequalities. In particular, wrt. Lemma 16, generation of this expert knowledge is straightforward. The corresponding inequality amounts to a line in the expert knowledge matrix A . Wrt. monotonicity we have experimented with a dedicated size analysis based on a simple static analysis of the given AST, as well as exploitation of the type annotations directly. For the latter, note that the coefficients in the basic potential functions $p_{((a_1, \dots, a_m), b)}$ are reflected in the corresponding type annotations. Hence, comparison of these (unknown) coefficients allows a sufficient size comparison of the data structures (i.e. trees) used in the program at hand.

For now, our exploitation of the expert knowledge is restricted to the monotonicity of logarithm together with the simple mathematical fact about logarithms presented in Lemma 16. To improve the efficiency and effectivity of the methodology, the following additions could be explored: (i) additional mathematical facts on the logarithm function; (ii) improvement of the dedicated size analysis supporting the applicability of monotonicity laws; (iii) incorporation of basic static analysis techniques, like the result of a reachability analysis, etc.

9. Related Work

To the best of our knowledge, the established type-and-effect system for the analysis of logarithmic amortised complexity is novel and also the semi-automated resource analysis of self-balanced data structures like splay trees is unparalleled in the literature. However, there is a vast amount of literature on (automated) resource analysis. Without hoping for completeness, we briefly mention Albert *et al.* (2008); Alias *et al.* (2010); Blanc *et al.* (2010); Gulwani and Zuleger (2010); Albert *et al.*

(2011); Alonso-Blas and Genaim (2012); Hermenegildo et al. (2012); Avanzini et al. (2015, 2016); Flores-Montoya (2017); Giesl et al. (2017) for an overview of the field.

(Constant) amortised cost analysis has been in particular pioneered by Martin Hofmann and his collaborators. Starting with seminal work on the static prediction of heap space usage (Hofmann and Jost, 2003; Hofmann and Rodriguez, 2013), the approach has been generalised to (lazy) functional programming (Jost et al., 2009, 2010; Hoffmann et al., 2011, 2012b,a) and rewriting (Hofmann and Moser, 2014, 2015). Automation of amortised resource analysis has also been greatly influenced by Hofmann, yielding to sophisticated tools for the analysis of high-order functional programs (Hoffmann and Hofmann, 2010b,a; Hoffmann, 2011), as well as of object-oriented programs (Hofmann and Rodriguez, 2013; Bauer et al., 2018). We mention here the highly sophisticated analysis behind the RaML prototype developed in Hoffmann and Sho (2014); Hoffmann and Shao (2015a,b); Hoffmann et al. (2017) and the RAJA tool (Hofmann and Rodriguez, 2013).

We now overview alternatives to conducting amortised cost analysis through the means of a type-and-effect system. The line of work Zuleger et al. (2011); Sinn et al. (2014, 2015, 2017); Fiedor et al. (2018) has focused on identifying abstractions resp. abstract program models that can be used for the automated resource analysis of imperative programs. The goal has been to identify program models that are sufficiently rich to support the inference of precise bounds and sufficiently abstract to allow for a scalable analysis, employing the size-change abstraction (Zuleger et al., 2011), (lossy) vector-addition systems (Sinn et al., 2014) and difference-constraint systems (Sinn et al., 2015, 2017). This work has led to the development of the tool LOOPUS, which performs amortised analysis for a class of programs that cannot be handled by related tools from the literature. Interestingly, LOOPUS infers worst-case costs from lexicographic ranking functions using arguments that implicitly achieve an amortised analysis (for details, we refer the reader to Sinn et al. (2017)). Another line of work has targeted the resource bound analysis of imperative and object-oriented programs through the extraction of recurrence relations from the program under analysis, whose closed-form solutions then allow one to infer upper bounds on resource usage (Albert et al., 2008, 2011; Alonso-Blas and Genaim, 2012; Flores-Montoya, 2017). Amortised analysis with recurrence relations has been discussed for the tools COSTA (AlonsoBlas and Genaim, 2012) and CoFloCo (Flores-Montoya, 2017). Amortised analysis has also been employed in the resource analysis for rewriting (Moser and Schneckeneither, 2020) and non-strict function programs, in particular, if *lazy evaluation* is conceived, cf. Jost et al. (2017).

Sublinear bounds are typically not in the focus of these tools, but can be inferred by some tools. In the recurrence relations-based approach to cost analysis (Albert et al., 2008, 2011) refinements of linear ranking functions are combined with criteria for divide-and-conquer patterns; this allows their tool PUBS to recognise logarithmic bounds for some problems, but examples such as *mergesort* or *splaying* are beyond the scope of this approach. Logarithmic and exponential terms are integrated into the synthesis of ranking functions in Chatterjee et al. (2017), making use of an insightful adaption of Farkas' and Handelman's lemmas. The approach is able to handle examples such as *mergesort*, but again not suitable to handle self-balancing data structures. A type-based approach to cost analysis for an ML-like language is presented in Wang et al. (2017), which uses the Master Theorem to handle divide-and-conquer-like recurrences. Very recently, support for the Master Theorem was also integrated for the analysis of rewriting systems by Winkler and Moser (2020), extending the work of Avanzini and Moser (2016) on the modular resource analysis of rewriting to so-called logically constrained rewriting systems (Fuhs et al., 2017). The resulting approach also supports the fully automated analysis of *mergesort*.

We also mention the quest for abstract program models whose resource bound analysis problem is decidable, and for which the obtainable resource bounds can be precisely characterised. We list here the size-change abstraction, whose worst-case complexity has been completely characterised as polynomial (with rational coefficients) (Colcombet et al., 2014; Zuleger, 2015), vector-addition systems (Brázdil et al., 2018; Zuleger, 2020), for which polynomial complexity can

be decided, and LOOP programs (Ben-Amram and Hamilton, 2019), for which multivariate polynomial bounds can be computed. We are not aware of similar results for program models that induce logarithmic bounds.

10. Conclusion

We have presented a novel amortised resource analysis based on the potential method. The method is rendered in a type-and-effect system. Our type system has been designed with the goal of automation. The novelty of our contribution is that this is the first approach towards an automation of a *logarithmic* amortised complexity analysis. In particular, we show how the precise logarithmic amortised cost of *splaying*—a central operation of Sleator and Tarjan’s splay trees—can be checked semi-automatically in our system. As our potential functions are logarithmic, we cannot directly encode the comparison between logarithmic expressions within the theory of linear arithmetic. This however is vital for, for example, expressing Schoenmakers’ and Nipkow’s (manual) analysis (Nipkow, 2015; Schoenmakers, 1993) in our type-and-effect system. To overcome this algorithmic challenge, we proposed several ideas for the *linearisation* of the induced constraint satisfaction problem. These efforts can be readily extended by expanding upon the *expert knowledge* currently employed, for example, via incorporation of the results of a static analysis performed in a pre-processing step. In future work, we aim at extension of the developed prototypes to a fully automated analysis of logarithmic amortised complexity. Here it may be profitable to expand the class of potential functions to take *linear* potential functions into account. This does not invalidate our soundness theorem.

In Memoriam.

Martin Hofmann and the fourth author have discussed and developed a large part of the theoretical body of this work together. Unfortunately, Martin’s tragic hiking accident in January 2018 prevented the conclusion of this collaboration. Due to Martin’s great interest and contributions to this work, we felt it fitting to include him as first author. We have tried our best to finalise the common conceptions and ideas. Still automation and continued research clarified a number of issues and also brought a different focus on various matters of the material presented. Martin Hofmann’s work was revolutionary in a vast amount of fields and it will continue to inspire future researchers—like he inspired us.

References

- Albert, E., Arenas, P., Genaim, S. and Puebla, G. (2008). Automatic inference of upper bounds for recurrence relations in cost analysis. In: *Proceedings of the 15th SAS*, vol. 5079, 221–237.
- Albert, E., Arenas, P., Genaim, S. and Puebla, G. (2011). Closed-form upper bounds in static cost analysis. *Journal of Automated Reasoning* 46 (2).
- Alias, C., Darte, A., Feautrier, P. and Gonnord, L. (2010). Multi-dimensional rankings, program termination, and complexity bounds of flowchart programs. In: *Proceedings of the 17th SAS*, vol. 6337, LNCS, 117–133.
- Alonso-Blas, D. E. and Genaim, S. (2012). On the limits of the classical approach to cost analysis. In: Miné, A. and Schmidt, D. (eds.) *Proceedings of the 19th SAS*, vol. 7460, LNCS, Springer, 405–421.
- Avanzini, M., Eguchi, N. and Moser, G. (2011). A path order for rewrite systems that compute exponential time functions. In: *Proceedings of the 22nd RTA*, vol. 10. LIPIcs, 123–138.
- Avanzini, M., Lago, U. D. and Moser, G. (2015). Analysing the complexity of functional programs: higher-order meets first-order. In: *Proceedings of the 20th ICFP*, ACM, 152–164.
- Avanzini, M. and Moser, G. (2016). A combination framework for complexity. *Information and Computation* 248 22–55.
- Avanzini, M., Moser, G. and Schaper, M. (2016). TcT: Tyrolean Complexity Tool. In: *Proceedings of the 22nd TACAS*, vol. 9636. LNCS, 407–423.
- Bauer, S., Jost, S. and Hofmann, M. (2018). Decidable inequalities over infinite trees. In: Barthe, G., Sutcliffe, G. and Veanes, M. (eds.), *Proceedings of the 22nd LPAR*, vol. 57. EPIc Series in Computing, EasyChair, 111–130.
- Ben-Amram, A. M. and Hamilton, G. W. (2019). Tight worst-case bounds for polynomial loop programs. In: Bojanczyk, M. and Simpson, A. (eds.), *Proceedings of the 22nd FOSSACS*, vol. 11425. LNCS, Springer, 80–97.

- Blanc, R., Henzinger, T. A., Hottelier, T. and Kovács, L. (2010). ABC: Algebraic bound computation for loops. In: *Proceedings of the 16th LPAR*, vol. 6355. LNCS, 103–118.
- Brázdil, T., Chatterjee, K., Kucera, A., Novotný, P., Velan, D. and Zuleger, F. (2018). Efficient algorithms for asymptotic bounds on termination time in VASS. In: Dawar, A. and Grädel, E. (eds.) *Proceedings of the 33rd LICS*, ACM, 185–194.
- Chatterjee, K., Fu, H. and Goharshady, A. K. (2017). Non-polynomial worst-case analysis of recursive programs. In: *Proceedings of the 29th CAV*, vol. 10427, LNCS, 41–63.
- Colcombet, T., Daviaud, L. and Zuleger, F. (2014). Size-change abstraction and max-plus automata. In: Csehaj-Varjú, E., Dietzfelbinger, M. and ésik, Z. (eds.) *Proceedings of 39th MFCS*, vol. 8634, LNCS, Springer, 208–219.
- Contejean, E., Marché, C., Tomás, A.-P. and Urbain, X. (2005). Mechanically proving termination using polynomial interpretations. *Journal of Automated Reasoning* 34 (4) 325–363.
- Fiedor, T., Holk, L., Rogalewicz, A., Sinn, M., Vojnar, T. and Zuleger, F. (2018). From shapes to amortized complexity. In: Dillig, I. and Palsberg, J. (eds.) *Proceedings of the 19th VMCAI*, vol. 10747, LNCS. Springer, 205–225.
- Flores-Montoya, A. (2017). *Cost Analysis of Programs Based on the Refinement of Cost Relations*. PhD thesis, Darmstadt University of Technology, Germany.
- Fuhs, C., Kop, C. and Nishida, N. (2017). Verifying procedural programs via constrained rewriting induction. *Transactions on Computational Logic* 18 (2) 14:1–14:50.
- Giesl, J., Aschermann, C., Brockschmidt, M., Emmes, F., Frohn, F., Fuhs, C., Hensel, J., Otto, C., Plücker, M., Schneider-Kamp, P., Ströder, T., Swiderski, S. and Thiemann, R. (2017). Analyzing program termination and complexity automatically with AProVE. *Journal of Automated Reasoning* 58 (1) 3–31.
- Gulwani, S. and Zuleger, F. (2010). The reachability-bound problem. In Zorn, B. G. and Aiken, A. (eds.) *PLDI*, ACM, 292–304.
- Hermenegildo, M., Bueno, F., Carro, M., López-Garca, P., Mera, E., Morales, J. and Puebla, G. (2012). An overview of ciao and its design philosophy. *Theory and Practice of Logic Programming* 12 (1–2) 219–252.
- Hoffmann, J. (2011). *Types with Potential: Polynomial Resource Bounds via Automatic Amortized Analysis*. PhD thesis, Ludwig-Maximilians-Universität München.
- Hoffmann, J., Aehlig, K. and Hofmann, M. (2011). Multivariate amortized resource analysis. In: *Proceedings of the 38th POPL*. ACM, 357–370.
- Hoffmann, J., Aehlig, K. and Hofmann, M. (2012a). Multivariate amortized resource analysis. *Transactions on Programming Languages and Systems* 34 (3) 14.
- Hoffmann, J., Aehlig, K. and Hofmann, M. (2012b). Resource aware ML. In *Proceedings of the 24th CAV*, vol. 7358, LNCS, 781–786.
- Hoffmann, J., Das, A. and Weng, S.-C. (2017). Towards automatic resource bound analysis for OCaml. In *Proceedings of the 44th POPL*. ACM, 359–373.
- Hoffmann, J. and Hofmann, M. (2010a). Amortized resource analysis with polymorphic recursion and partial big-step operational semantics. In: *Proceedings of the 8th APLAS*, vol. 6461. LNCS, 172–187.
- Hoffmann, J. and Hofmann, M. (2010b). Amortized resource analysis with polynomial potential. In: *Proceedings of the 19th ESOP*, vol. 6012, LNCS, 287–306.
- Hoffmann, J. and Shao, Z. (2015a). Automatic static cost analysis for parallel programs. In: *Proceedings of the 24th ESOP*, vol. 9032. LNCS, 132–157.
- Hoffmann, J. and Shao, Z. (2015b). Type-based amortized resource analysis with integers and arrays. *Journal of Functional Programming* 25.
- Hoffmann, J. and Sho, Z. (2014). Type-based amortized resource analysis with integers and arrays. In: *Proceedings of the 12th FLOPS*, vol. 8475. LNCS, 152–168.
- Hofmann, M. and Jost, S. (2003). Static prediction of heap space usage for first-order functional programs. In: *Proceedings of the 30th POPL*, ACM, 185–197.
- Hofmann, M. and Moser, G. (2014). Amortised resource analysis and typed polynomial interpretations. In: *Proceedings of Joint 25th RTA and 12th TLCA*, vol. 8560. LNCS, 272–286.
- Hofmann, M. and Moser, G. (2015). Multivariate amortised resource analysis for term rewrite systems. In: *Proceedings of the 13th TLCA*, vol. 38, *LIPICs*, 241–256.
- Hofmann, M. and Moser, G. (2018). Analysis of logarithmic amortised complexity.
- Hofmann, M. and Rodriguez, D. (2013). Automatic type inference for amortised heap-space analysis. In Felleisen, M. and Gardner, P. (eds.) *Proceedings of the 22nd ESOP*, vol. 7792. LNCS, Springer, 593–613.
- Jost, S., Hammond, K., Loidl, H.-W. and Hofmann, M. (2010). Static determination of quantitative resource usage for higher-order programs. In: *Proceedings of the 37th POPL*, ACM, 223–236.
- Jost, S., Loidl, H.-W., Hammond, K., Scaife, N. and Hofmann, M. (2009). “Carbon Credits” for resource-bounded computations using amortised analysis. In: *Proceedings of the 2nd FM*, vol. 5850, LNCS, 354–369.
- Jost, S., Vasconcelos, P., Florido, M. and Hammond, K. (2017). Type-based cost analysis for lazy functional languages. *Journal of Automated Reasoning* 59 (1) 87–120.
- Kahn, D. M. and Hoffmann, J. (2020). Exponential automatic amortized resource analysis. In Goubault-Larrecq, J. and König, B. (eds.) *Proceedings of the 23rd FOSSACS*, vol. 12077. LNCS, Springer, 359–380.

- Moser, G. and Schneckenreither, M. (2018). Automated amortised resource analysis for term rewrite systems. In: *Proceedings of the 14th FLOPS*, vol. 10818, LNCS, 214–229.
- Moser, G. and Schneckenreither, M. (2020). Automated amortised resource analysis for term rewrite systems. *Science of Computer Programming* 185.
- Nipkow, T. (2015). Amortized complexity verified. In: *Proceedings of the 6th ITP*, vol. 9236, LNCS, 310–324.
- Okasaki, C. (1999). *Purely Functional Data Structures*, Cambridge University Press.
- Pierce, B. (2002). *Types and Programming Languages*, MIT Press.
- Podelski, A. and Rybalchenko, A. (2004). A complete method for the synthesis of linear ranking functions. In: *Proceedings of the 5th VMCAI*, vol. 2937, LNCS, 239–251.
- Schoenmakers, B. (1993). A systematic analysis of splaying. *Information Processing Letters* 45 (1) 41–50.
- Schrijver, A. (1999). *Theory of Linear and Integer Programming*, Wiley.
- Sinn, M., Zuleger, F. and Veith, H. (2014). A simple and scalable static analysis for bound analysis and amortized complexity analysis. In: *Proceedings of the 26th CAV*, vol. 8559, LNCS, 745–761.
- Sinn, M., Zuleger, F. and Veith, H. (2015). Difference constraints: An adequate abstraction for complexity analysis of imperative programs. In: Kaivola, R. and Wahl, T. (eds.) *FMCAD*, IEEE, 144–151.
- Sinn, M., Zuleger, F. and Veith, H. (2017). Complexity and resource bound analysis of imperative programs using difference constraints. *Journal of Automated Reasoning* 59 (1) 3–45.
- Sleator, D. and Tarjan, R. (1985). Self-adjusting binary search trees. *Journal of Alternative and Complementary Medicine* 32 (3) 652–686.
- Tarjan, R. (1985). Amortized computational complexity. *SIAM Journal on Algebraic Discrete Methods* 6 (2) 306–318.
- Wang, P., Wang, D. and Chlipala, A. (2017). TiML: A functional language for practical complexity analysis with invariants. *Proceedings of the ACM on Programming Languages* 1 (OOPSLA).
- Winkler, S. and Moser, G. (2020). Runtime complexity analysis of logically constrained rewriting. In: *Proceedings of the LOPSTR 2020*.
- Zuleger, F. (2015). Asymptotically precise ranking functions for deterministic size-change systems. In: Beklemishev, L. D. and Musatov, D. V. (eds.) *Proceedings of 10th CSR*, vol. 9139, LNCS, Springer, 426–442.
- Zuleger, F. (2020). The polynomial complexity of vector addition systems with states. In: Goubault-Larrecq, J. and König, B. (eds.) *Proceedings of the 23rd FOSSACS*, vol. 12077, LNCS, Springer, 622–641.
- Zuleger, F., Gulwani, S., Sinn, M. and Veith, H. (2011). Bound analysis of imperative programs with the size-change abstraction. In: Yahav, E. (ed.) *Proceedings of the 18th SAS*, vol. 6887, LNCS, Springer, 280–297.