Amortized Resource Analysis with Polymorphic Recursion and Partial Big-Step Operational Semantics

Extended Version

Jan Hoffmann* and Martin Hofmann

Ludwig-Maximilians-Universität München

Abstract. This paper studies the problem of statically determining upper bounds on the resource consumption of first-order functional programs. A previous work approached the problem with an automatic type-based amortized analysis for polynomial resource bounds. The analysis is parametric in the resource and can be instantiated to heap space, stack space, or clock cycles. Experiments with a prototype implementation have shown that programs are analyzed efficiently and that the computed bounds exactly match the measured worst-case resource behavior for many functions. This paper describes the inference algorithm that is used in the implementation of the system. It can deal with resource-polymorphic recursion which is required in the type derivation of many functions. The computation of the bounds is fully automatic if a maximal degree of the polynomials is given. The soundness of the inference is proved with respect to a novel operational semantics for partial evaluations to show that the inferred bounds hold for terminating as well as non-terminating computations. A corollary is that run-time bounds also establish the termination of programs.

1 Introduction

The quantitative analysis of algorithms is a classic problem in computer science. For many applications in software development it is necessary to obtain not only asymptotic bounds but rather specific upper bounds for concrete implementations. This is especially the case for the development of embedded and safety-critical systems.

Even for basic programs, manual analysis of the specific (non-asympt.) costs is tedious and error-prone. The problem gets increasingly complex for high-level programming languages, since one needs to be aware of the translation performed by the compiler. As a result, automatic methods for analyzing the resource behavior of programs have been the subject of extensive research (see §8).

Our approach to the problem follows a line of research that was initiated by Hofmann and Jost [1]. It is based on the potential method of amortized

* Supported by the DFG Graduiertenkolleg 1480 (PUMA).

analysis that has been invented by Sleator and Tarjan [2] to simplify the manual reasoning about the costs of a sequence of operations that manipulate a data structure. [1] showed that a fully automatic amortized resource analysis can efficiently compute bounds on the heap-space consumption of many (first-order) functional programs that admit *linear* resource bounds. The limitation to linear bounds and accordingly linear constraints was essential for the efficiency of the analysis. Subsequent research considerably extended the range of type-based amortized analysis, but the restriction to linear bounds remained. Examples are the extensions of type-based amortized analysis to object-oriented programs [3, 4], to generic resource metrics [5, 6], to polymorphic and higher-order programs [7], and to Java-like bytecode by means of separation logic [8].

Somewhat unexpectedly, we recently discovered a technique [9] that yields an automatic amortized analysis for polynomial bounds while still relying on linear constraint solving only. The resulting system efficiently computes resource bounds for first-order functional programs that admit bounds that are sums $\sum p_i(n_i)$ of univariate polynomials p_i. This includes bounds on the heap-space usage and the number of evaluation steps for a number of interesting functions such as quick sort, merge sort, insertion sort, longest common subsequence via dynamic programming, breadth-first traversal of a tree using a functional queue, and sieve of Eratosthenes.

The system has been implemented for *Resource Aware ML (RAML)* which is a first-order fragment of OCAML. It is available online[1] and can be run directly in a web browser to analyze example programs and user-generated code. Our experiments show that the computed bounds exactly match the measured worst-case behavior in many cases (see Table 1). For example we obtain tight evaluation-step bounds for quick sort and insertion sort (see §7).

The basic idea of the analysis is to fix a maximal degree k and then to collect linear constraints on the coefficients of polynomials of this degree. One can iteratively increase the degree so as to avoid costly computations earlier on. A fine point arises from the fact that polynomials must be nonnegative and monotone and that in order for allowing local constraint generation for pattern matches the class of allowed polynomials must be closed under the operation $p(n) \mapsto p(n+1)-p(n)$. This naturally leads to nonnegative linear combinations of binomial coefficients. Other fine points are the treatment of sharing and aliasing.

A further challenge for the inference of polynomial bounds is the need to deal with *resource-polymorphic recursion* (see §2), which is required to type most of the above example programs. However, it seems to be a hard problem to infer general resource polymorphic recursion even for the original linear system.

In this paper we present a pragmatic approach to resource-polymorphic recursion that works well and efficiently in practice. It infers types for most functions that admit a type-derivation, including the above examples. Nevertheless, it is not complete with respect to resource-polymorphic typing rules. §6 contains a somewhat artificial function with a linear heap-space consumption that admits

[1] See `http://raml.tcs.ifi.lmu.de`.

a resource-polymorphic typing that can neither be inferred in our system nor in the linear system [1].

The main theorem of the paper (see §5) shows that the resource bounds are sound with respect to a big-step operational semantics. A dissatisfying feature of classical big-step semantics is that it does not provide evaluation judgments for non-terminating evaluations. As a result, the soundness theorems for amortized resource analyses have in the past been formulated for terminating evaluations only [1, 5, 7].

A secondary contribution of this paper is the introduction of a novel big-step operational semantics for partial evaluations which agrees with the usual big-step semantics on terminating computations. In this way, we retain the advantages of big-step semantics (shorter, less syntactic proofs; better agreement (arguably) with actual behaviour of computers) while capturing the resource behaviour of non-terminating programs. This enables the proof of an improved soundness result: if the type analysis has established a resource bound for an expression then the resource consumption of its (possibly non-terminating) evaluation does not exceed the bound. It follows that run-time bounds also ensure termination.

This paper complements a previous paper [9]. The contributions we make here are as follows.

- We introduce a novel operational semantics for partial evaluations that allows a simplified and improved soundness theorem (in §4).
- We present algorithmic typing rules that are used by the inference algorithm (in §5).
- An extended soundness proof shows that the inferred bounds hold for both terminating and non-terminating computations (Thm. 5).
- We describe a new inference algorithm that efficiently computes resource-polymorphic types for most functions for which such a type exists (in §6).

2 Informal Presentation

Linear Potential The general idea of type-based amortized analysis for functional programs has been introduced in [1] as follows. First, inductive data structures are statically annotated with a positive rational number q to define a non-negative potential $\Phi(n) = q \cdot n$ as a function of the size n of the data. Second, the potential is shown to be sufficient to pay for all operations that are performed on this data structure during any possible evaluation of the program. The initial potential (summed over all input data) then describes an upper bound on the resource costs. We illustrate the idea by analyzing the heap-space consumption of the function *attach* below.

attach(x,l) = **match** l **with** | nil → nil | (y::ys) → (x,y)::(attach(x,ys))

It takes an integer and a list of integers and returns a list of pairs of integers in which the first argument is paired with each element of the list. If we assume that a list element for a pair of integers has size 3 (two cells to store the integers and one for the pointer to the next element) then the heap-space usage of an evaluation of *attach(*x*,l)* is $3|l|$ memory cells.

In order to infer an upper bound on the heap-space usage of the function we annotate the type of *attach* with a priori unknown resource-annotations s, s', q and p that range over non-negative rational numbers. The intuitive meaning of the resulting type *attach:* $(int, L^q(int)) \xrightarrow{s/s'} L^p(int, int)$ is as follows: to evaluate *attach(*x*,l)* one needs q memory cells per element in the list l and s additional memory cells. After the evaluation there are s' memory cells and p cells per element of the returned list left. We say that the list l has potential $\Phi(l, q) = q \cdot |l|$ and that *l' = attach(*x*,l)* has potential $\Phi(l', p) = p \cdot |l'|$.

The problem of computing a resource bound then amounts to finding valid instantiations of the resource variables, i.e., a potential that suffices to cover the costs of any possible evaluation. The validity of an instantiation can be verified statically in a sound albeit not complete type-based analysis of the program text. A valid resource annotation for *attach* can be obtained by setting $q = 3$ and $s = s' = p = 0$. The computed upper bound on the heap-space costs is then $3n$ where n is length of the input list. Another possible instantiation would be $q = 6$, $p = 3$, and $s = s' = 0$. The resulting typing of *attach* could be used for the inner occurrence of *attach* to type an expression like *attach(x,attach(z,ys))*. The associated upper bound on the heap-space costs for the evaluation of the expression is then $6|ys|$.

The use of linear potential functions relieves one of the burden of having to manipulate symbolic expressions during the analysis by a priori fixing their format. This gives rise to a particularly efficient inference algorithm for the type annotations. It works like a standard constraint-based type inference in which simple linear constraints are collected as each type rule is applied. The constraints are then solved by linear programming. To see the basic idea, consider the function *attach* in which expressions of type list are annotated with variables $q, p, r, \ldots$ that range over $\mathbb{Q}^+$. The intended meaning of l^q is that l is of type $L^q(A)$ for some type A.

$$\begin{array}{ll} \mathrm{attach(x,l}^{q}) = \textbf{match}\ \mathrm{l}^{q'}\ \textbf{with} & |\ \mathrm{nil} \to \mathrm{nil}^{p} \\ & |\ \ (\mathrm{y::ys}^{r}) \to ((\mathrm{x,y})\mathrm{::}(\mathrm{attach\ (x,ys}^{q}))^{p})^{p} \end{array}$$

The syntax-directed inference then computes inequalities like $q' + s \geq 3 + p + s$. It expresses the fact that the potential q' of the first list element and the initial potential s need to cover the costs for the cons operation (3 memory cells), the potential p of a list element of the result, and the input potential s of the recursive call. To pay the cost during the recursion we require the annotation of the function arguments and the result of the recursive call to match their specification ($s = q$ and $t = p$ in the case of *attach*). The function is then used *resource-monomorphically*, i.e., with the same annotations as in the result and the arguments of the call.

Polynomial Potential Our previous work [9] showed that an automatic amortized analysis can also be used to derive *polynomial* resource bounds by extracting *linear* inequalities from a program. The main innovation is the use of potential-functions of the form $\sum_{i=1,\ldots,k} q_i\binom{n}{i}$ with $q_i \geq 0$. They are attached to inductive data structures via type annotations of the form $\vec{q} = (q_1, \ldots, q_k)$ with $q_i \in \mathbb{Q}^+$. For instance, the typing $l{:}L^{(3,2,1)}(\mathit{int})$, defines the potential $\Phi(l,(3,2,1)) = 3|l| + 2\binom{|l|}{2} + 1\binom{|l|}{3}$. One intuition for these numbers is as follows: The annotation $\vec{q}$ assigns the potential q_1 to every element of the list, the potential q_2 to every element of every proper suffix of the list, q_3 to the elements of the suffixes of the suffixes, etc.

The use of binomial coefficients rather than powers of variables has many advantages as discussed in [9]. In particular, the identity $\sum_{i=1,\ldots,k} q_i\binom{n+1}{i} = q_1 + \sum_{i=1,\ldots,k-1} q_{i+1}\binom{n}{i} + \sum_{i=1,\ldots,k} q_i\binom{n}{i}$ gives rise to a local typing rule for *cons match* which naturally allows the typing of both recursive calls and other calls to subordinate functions in branches of a pattern match.

This identity forms the mathematical basis of the *additive shift* $\lhd$ of a type annotation which is defined by $\lhd(q_1, \ldots, q_k) = (q_1 + q_2, \ldots, q_{k-1} + q_k, q_k)$. For example, it appears in the typing $\mathit{tail}{:}L^{\vec{q}}(\mathit{int}) \xrightarrow{0/q_1} L^{\lhd(\vec{q})}(\mathit{int})$ of the function *tail* that removes the first element from a list. The idea underlying the additive shift is that the potential resulting from the contraction $\mathit{xs}{:}L^{\lhd(\vec{q})}(\mathit{int})$ of a list $(x{::}xs){:}L^{\vec{q}}(\mathit{int})$ (usually in a pattern match) is used for three purposes: (i) to pay the constant costs after and before the recursive calls (q_1), (ii) to fund calls to auxiliary functions ($(q_2, \ldots, q_n)$), and (iii) to pay for the recursive calls ($(q_1, \ldots, q_n)$).

To see how the polynomial potential annotations are used to compute polynomial resource bounds, consider the function *pairs* that computes the two-element subsets of a given set (representing sets as tuples or lists).

pairs l = **match** l **with** | nil → nil | (x::xs) → append(attach(x,xs),pairs xs)

The function *append* consumes 3 memory cells for every element in the first argument. Similar to *attach* we can compute a tight resource bound for *append* by inferring the type *append:* $(L^{(3)}(\mathit{int},\mathit{int}), L^{(0)}(\mathit{int},\mathit{int})) \xrightarrow{0/0} L^{(0)}(\mathit{int},\mathit{int})$.

The evaluation of the expression *pairs(l)* consumes 6 memory cells per element of every sub-list (suffix) of l. The type that our system infers for *pairs* is $L^{(0,6)}(\mathit{int}) \xrightarrow{0/0} L^{(0)}(\mathit{int},\mathit{int})$. It states that a list l in an expression *pairs(l)* has the potential $\Phi(l,(0,6)) = 0 \cdot |l| + 6 \cdot \binom{|l|}{2}$ and thus furnishes a tight upper bound on the heap-space usage. To type the functions body, the additive shift assigns the type $\mathit{xs}{:}L^{(0+6,6)}(\mathit{int})$ to the variable *xs* in the pattern match. The potential is shared between the two occurrences of *xs* in the following expression by using $\mathit{xs}{:}L^{(6,0)}(\mathit{int})$ to pay for *append* and *attach* (ii) and using $\mathit{xs}{:}L^{(0,6)}(\mathit{int})$ to pay for the recursive call of *pairs* (iii); the constant costs (i) are zero in this example.

To compute the bound, we start with an annotation of the list types with resource variables as before.

pairs l = **match** $\mathrm{l}^{(q_1,q_2)}$ **with** | nil → nil
| (x :: $\mathrm{xs}^{(p_1,p_2)}$) → append(attach(x,$\mathrm{xs}^{(r_1,r_2)}$),pairs $\mathrm{xs}^{(s_1,s_2)}$)

The constraints that our type system computes include $q_2{\geq}p_2$ and $q_1{+}q_2{\geq}p_1$ (additive shift); $p_1{=}r_1{+}s_1$ and $p_2{=}r_2{+}s_2$ (sharing between two variables); $r_1{\geq}6$

(pay for non-recursive function calls); $q_1{=}s_1$, $q_2{=}s_2$ (pay for the recursive call). This system is solvable by $q_2 = s_2 = p_1 = p_2 = r_1 = 6$ and $q_1 = s_1 = r_2 = 0$.

Polymorphic Recursion As in the linear case, we require in the constraint system that the type of the recursive call of *pairs* matches its specification ($q_i = s_i$). But other than in the linear case, such a resource-monomorphic approach results in an unsolvable linear program for many non-tail-recursive functions with a super linear resource behavior. We illustrate this with the function *pairs'* that is a modification of *pairs* in which we permute the arguments of *append* and hence replace the expression in the cons-branch of the pattern match with *append(pairs' xs,attach(x,xs))*. The heap-space usage of *pairs'* is $3\binom{n}{2}+3\binom{n}{3}$ since *append* is called with the intermediate results of *pairs'* in the first argument and thus consumes $\sum_{2\leq i<n}\binom{i}{2} = \binom{n}{3}$ memory cells.

The resource-polymorphic system determines an exact heap-space bound for the function *pairs'* by computing the typing $L^{(0,3,3)}(\mathit{int})\xrightarrow{0/0}L^{(0)}(\mathit{int},\mathit{int})$. Similar to the case of *pairs* the additive shift assigns the type $L^{(3,6,3)}(\mathit{int})$ to *xs* in the cons-branch. The linear potential $\mathit{xs}{:}L^{(3,0,0)}(\mathit{int})$ is passed on to the occurrence of *xs* in *attach*. But in order to pay the costs of *append* we have to assign a linear potential to the result of the recursive call and thus use the alternate typing *pairs':* $L^{(0,6,3)}(\mathit{int})\xrightarrow{0/0}L^{(3)}(\mathit{int},\mathit{int})$. The need of passing on potential of degree at most $k-1$ to the output of a function with a resource consumption of degree k is quite common in typical functions. It is present in the derivation of time bounds for most non-tail-recursive functions that we considered, e.g., quick sort and insertion sort. The classic (resource-monomorphic) inference approach of requiring the type of the recursive call to match its specification fails for these functions and it was a non-trivial problem to address it with an efficient solution.

Cost-Free Resource Metric Our pragmatic approach is to introduce a special *cost-free* resource metric that assigns zero costs to every evaluation step. A cost-free function type *f:* $A\xrightarrow{a/a'}B$ then describes how to pass potential from x to $f(x)$ without paying for resource usage. Any concrete typing for a given resource metric can be superposed with a *cost-free* typing to obtain another typing for the given resource metric (cf. the solution of inhomogeneous systems by superposition with homogeneous solutions in linear algebra).

We illustrate the idea using *pairs'* again. First, we derive the cost-free types *attach:* $(\mathit{int},L^{(3)}(\mathit{int}))\xrightarrow{0/0}L^{(3)}(A)$ and *append:* $(L^{(3)}(A),L^{(3)}(A))\xrightarrow{0/0}L^{(3)}(A)$ for $A = (\mathit{int},\mathit{int})$. The type inference for, e.g., *attach* works as outlined above with the inequality $q' + s \geq 3 + p + s$ replaced with $q' + s \geq p + s$. Similar, we can assign *pairs'* the cost-free type $L^{(0,3)}(\mathit{int})\xrightarrow{0/0}L^{(3)}(\mathit{int},\mathit{int})$. The typing $\mathit{xs}{:}L^{(3,3)}(\mathit{int})$ that results from the additive shift is then used as $\mathit{xs}{:}L^{(3,0)}(\mathit{int})$ in *attach* and as $\mathit{xs}{:}L^{(0,3)}(\mathit{int})$ in the recursive call.

If we now want to infer the type of a function with respect to some cost metric then we deal with recursive calls by requiring them to match the functions type specification and to optionally pass potential to the result via a cost-free type. The cost-free type is then inferred resource-monomorphically. In the case of the heap-space consumption of *pairs'* we would first infer that the recursive call has to be of the form $L^{(0+q_1,3+q_2,3)}(\mathit{int}){\to}L^{(0+p_1)}(\mathit{int},\mathit{int})$, where $L^{(q_1,q_2)}(\mathit{int}){\to}L^{(p_1)}(\mathit{int},\mathit{int})$ is a cost-free type. We then infer like in the linear case that $q_1 = 0$ and $q_2 = p_1 = 3$.

This method cannot infer every resource-polymorphic typing with respect to declarative type derivations with polymorphic recursion. This would mean to start with a (possibly infinite) set of annotated types for each function and to justify each function type with a type derivation that uses types from the initial set. With respect to this declarative view, the inference algorithm in this paper can compute every set of types for a function f that has the form $\Sigma(f) = \{T + q \cdot T_i \mid q \in \mathbb{Q}^+, 1 \leq i \leq m\}$ for a resource-annotated function type T, cost-free function types T_i, and m recursive calls of f in its function body. Since many resource-polymorphic type derivations feature a set of function types of this format, our approach leads to an effective inference method.

In the algorithmic type rules (Fig. 3) we directly integrated the above format of $\Sigma(f)$ in the rule T:FunApp for function applications to enable an efficient inference. We prove the soundness theorem (Thm. 5) for the algorithmic rules since we do not present declarative type rules in this paper.

3 Resource Aware ML

RAML (Resource Aware ML) is a first-order functional language with ML-style syntax, booleans, integers, pairs, lists, recursion and pattern match. We deliberately focus on a simple language to describe the type inference here. However, it has been shown that type-based amortized analysis for linear bounds can deal with more advanced language features. This includes polymorphism [7], higher-order programs [7], destructive pattern matches [5, 1] and inductive data types [5, 7].

In the implementation of RAML we included a destructive pattern match and the extended version of [9] describes how polynomial potential can be applied to tree-like data types. The inference algorithm can easily be adopted to handle these extensions.

To simplify typing rules and semantics in this paper, we define the following *expressions of RAML* to be in *let normal form*. In the implementation we directly work with unrestricted expressions.

$$\begin{aligned}
e ::=\; & () \mid \mathit{True} \mid \mathit{False} \mid n \mid x \mid x_1 \;\mathit{binop}\; x_2 \mid f(x_1,\ldots,x_n) \mid \mathit{let}\; x = e_1 \;\mathit{in}\; e_2 \\
& \mid \mathit{if}\; x \;\mathit{then}\; e_t \;\mathit{else}\; e_f \mid (x_1,x_2) \mid \mathit{match}\; x \;\mathit{with}\; (x_1,x_2) \to e \\
& \mid \mathit{nil} \mid \mathit{cons}(x_h,x_t) \mid \mathit{match}\; x \;\mathit{with}\; \mid \mathit{nil} \to e_1 \mid \mathit{cons}(x_h,x_t) \to e_2 \\
\mathit{binop} ::=\; & + \mid - \mid * \mid \mathit{mod} \mid \mathit{div} \mid \mathit{and} \mid \mathit{or}
\end{aligned}$$

We define the well-typed expressions of RAML by assigning a *simple type*, a usual ML type without resource annotations, to well-typed expressions. Simple types are data types and first-order types as given by the grammars below.

$$A ::= \mathit{unit} \mid \mathit{bool} \mid \mathit{int} \mid L(A) \mid (A,A) \qquad\qquad F ::= A \to A$$

A *typing context* Γ is a partial, finite mapping from variable identifiers to data types. A *signature* Σ is a finite, partial mapping of function identifiers to first-order types. The typing judgment $\Gamma \vdash_{\Sigma} e : A$ states that the expression e has type A under the signature Σ in the context Γ. The typing rules that define the typing judgment are standard and identical with the resource-annotated typing rules from §5 if the resource annotations are omitted.

Each *RAML program* consists of a signature Σ and a family $(e_f, y_f)_{f \in \mathrm{dom}(\Sigma)}$ of expressions with a distinguished variable identifier such that $y_f{:}A \vdash_{\Sigma} e_f{:}B$ if $\Sigma(f) = A \to B$.

4 Operational Semantics

In the following we define a big-step operational semantics that measures the quantitative resource consumption of programs. It is parametric in the resource of interest and can measure every quantity whose usage in a single evaluation step can be bounded by a constant. The actual constants for a step on a specific system architecture can be derived by analyzing the translation of the step in the compiler implementation for that architecture [5].

The semantics is formulated with respect to a stack and a heap as usual: A value $v \in \mathrm{Val}$ is either a location $l \in \mathrm{Loc}$, a boolean constant b, an integer n, a null value Null or a pair of values (v_1, v_2). A *heap* is a finite partial mapping $\mathcal{H} : \mathrm{Loc} \to \mathrm{Val}$ that maps locations to values. A *stack* is a finite partial mapping $\mathcal{V} : \mathrm{VID} \to \mathrm{Val}$ from variable identifiers to values.

Since we also consider resources like memory that can become available during an evaluation, we have to track the *watermark* of the resource usage, i.e., the maximal number of resources units that are simultaneously used during an evaluation. In order to derive a watermark of a sequence of evaluations from the watermarks of the sub evaluations one has also to take into account the number of resource units that are available after each sub evaluation.

The operational evaluation rules in Fig. 1 thus define an evaluation judgment of the form $\mathcal{V}, \mathcal{H} \vdash e \rightsquigarrow v, \mathcal{H}' \mid (q, q')$ expressing the following. If the stack $\mathcal{V}$ and the initial heap $\mathcal{H}$ are given then the expression e evaluates to the value v and the new heap $\mathcal{H}'$. In order to evaluate e one needs at least $q \in \mathbb{Q}^+$ resource units and after the evaluation there are at least $q' \in \mathbb{Q}^+$ resource units available. The actual resource consumption is then $\delta = q - q'$. The quantity δ is negative if resources become available during the execution of e.

In contrast to similar versions in earlier works there is at most one pair (q, q') such that $\mathcal{V}, \mathcal{H} \vdash e \rightsquigarrow v, \mathcal{H}' \mid (q, q')$ for a given expression e, a heap $\mathcal{H}$ and a stack $\mathcal{V}$. The non-negative number q is the watermark of resources that are used simultaneously during the evaluation.

It is handy to view the pairs (q, q') in the evaluation judgments as elements of a monoid[2] $\mathcal{R} = (\mathbb{Q}^+ \times \mathbb{Q}^+, \cdot)$. The neutral element is $(0, 0)$ which means

[2] In fact, it is possible to define the evaluation more abstractly with respect to an arbitrary monoid M.

$$\frac{}{\mathcal{V},\mathcal{H} \vdash () \rightsquigarrow \textsc{Null}, \mathcal{H} \mid K^{\mathrm{unit}}}\ \textsc{E:ConstU} \qquad \frac{b \in \{\mathit{True}, \mathit{False}\}}{\mathcal{V},\mathcal{H} \vdash b \rightsquigarrow b, \mathcal{H} \mid K^{\mathrm{bool}}}\ \textsc{E:ConstB}$$

$$\frac{n \in \mathbb{Z}}{\mathcal{V},\mathcal{H} \vdash n \rightsquigarrow n, \mathcal{H} \mid K^{\mathrm{int}}}\ \textsc{E:ConstI} \qquad \frac{x_1, x_2 \in \mathrm{dom}(\mathcal{V}) \qquad v = \mathit{op}(\mathcal{V}(x_1), \mathcal{V}(x_2))}{\mathcal{V},\mathcal{H} \vdash x_1\ \mathit{op}\ x_2 \rightsquigarrow v, \mathcal{H} \mid K^{\mathrm{op}}}\ \textsc{E:BinOp}$$

$$\frac{x \in \mathrm{dom}(\mathcal{V})}{\mathcal{V},\mathcal{H} \vdash x \rightsquigarrow \mathcal{V}(x), \mathcal{H} \mid K^{\mathrm{var}}}\ \textsc{E:Var} \qquad \frac{x_1, x_2 \in \mathrm{dom}(\mathcal{V}) \qquad v = (\mathcal{V}(x_1), \mathcal{V}(x_2))}{\mathcal{V},\mathcal{H} \vdash (x_1, x_2) \rightsquigarrow v, \mathcal{H} \mid K^{\mathrm{pair}}}\ \textsc{E:Pair}$$

$$\frac{\mathcal{V}(x) = v \qquad [y_f \mapsto v], \mathcal{H} \vdash e_f \rightsquigarrow v', \mathcal{H}' \mid (q, q')}{\mathcal{V},\mathcal{H} \vdash f(x) \rightsquigarrow v', \mathcal{H}' \mid K_1^{\mathrm{app}} \cdot (q, q') \cdot K_2^{\mathrm{app}}}\ \textsc{E:FunApp}$$

$$\frac{\mathcal{V},\mathcal{H} \vdash e_1 \rightsquigarrow v_1, \mathcal{H}_1 \mid (q, q') \qquad \mathcal{V}[x \mapsto v_1], \mathcal{H}_1 \vdash e_2 \rightsquigarrow v_2, \mathcal{H}_2 \mid (p, p')}{\mathcal{V},\mathcal{H} \vdash \mathit{let}\ x = e_1\ \mathit{in}\ e_2 \rightsquigarrow v_2, \mathcal{H}_2 \mid K_1^{\mathrm{let}} \cdot (q, q') \cdot K_2^{\mathrm{let}} \cdot (p, p') \cdot K_3^{\mathrm{let}}}\ \textsc{E:Let}$$

$$\frac{\mathcal{V}(x) = \mathit{True} \qquad \mathcal{V},\mathcal{H} \vdash e_t \rightsquigarrow v, \mathcal{H}' \mid (q, q')}{\mathcal{V},\mathcal{H} \vdash \mathit{if}\ x\ \mathit{then}\ e_t\ \mathit{else}\ e_f \rightsquigarrow v, \mathcal{H}' \mid K_1^{\mathrm{conT}} \cdot (q, q') \cdot K_2^{\mathrm{conT}}}\ \textsc{E:CondT}$$

$$\frac{\mathcal{V}(x) = \mathit{False} \qquad \mathcal{V},\mathcal{H} \vdash e_f \rightsquigarrow v, \mathcal{H}' \mid (q, q')}{\mathcal{V},\mathcal{H} \vdash \mathit{if}\ x\ \mathit{then}\ e_t\ \mathit{else}\ e_f \rightsquigarrow v, \mathcal{H}' \mid K_1^{\mathrm{conF}} {\cdot} (q, q') {\cdot} K_2^{\mathrm{conF}}}\ \textsc{E:CondF}$$

$$\frac{\mathcal{V}(x) = (v_1, v_2) \qquad \mathcal{V}[x_1 \mapsto v_1, x_2 \mapsto v_2], \mathcal{H} \vdash e \rightsquigarrow v, \mathcal{H}' \mid (q, q')}{\mathcal{V},\mathcal{H} \vdash \mathit{match}\ x\ \mathit{with}\ (x_1, x_2) \to e \rightsquigarrow v, \mathcal{H}' \mid K_1^{\mathrm{matP}} \cdot (q, q') \cdot K_2^{\mathrm{matP}}}\ \textsc{E:MatP}$$

$$\frac{}{\mathcal{V},\mathcal{H} \vdash \mathit{nil} \rightsquigarrow \textsc{Null}, \mathcal{H} \mid K^{\mathrm{nil}}}\ \textsc{E:Nil} \qquad \frac{\begin{array}{c} x_h, x_t \in \mathrm{dom}(\mathcal{V}) \\ v = (\mathcal{V}(x_h), \mathcal{V}(x_t)) \qquad l \notin \mathrm{dom}(\mathcal{H}) \end{array}}{\mathcal{V},\mathcal{H} \vdash \mathit{cons}(x_h, x_t) \rightsquigarrow l, \mathcal{H}[l \mapsto v] \mid K^{\mathrm{cons}}}\ \textsc{E:Cons}$$

$$\frac{\mathcal{V}(x) = \textsc{Null} \qquad \mathcal{V},\mathcal{H} \vdash e_1 \rightsquigarrow v, \mathcal{H}' \mid (q, q')}{\begin{array}{c} \mathcal{V},\mathcal{H} \vdash \mathit{match}\ x\ \mathit{with}\ \mid \mathit{nil} \to e_1 \mid \mathit{cons}(x_h, x_t) \to e_2 \\ \rightsquigarrow v, \mathcal{H}' \mid K_1^{\mathrm{matN}} \cdot (q, q') \cdot K_2^{\mathrm{matN}} \end{array}}\ \textsc{E:MatN}$$

$$\frac{\mathcal{V}(x){=}l \qquad \mathcal{H}(l){=}(v_h, v_t) \qquad \mathcal{V}[x_h{\mapsto}v_h, x_t{\mapsto}v_t], \mathcal{H} \vdash e_2 \rightsquigarrow v, \mathcal{H}' \mid (q, q')}{\begin{array}{c} \mathcal{V},\mathcal{H} \vdash \mathit{match}\ x\ \mathit{with}\ \mid \mathit{nil} \to e_1 \mid \mathit{cons}(x_h, x_t) \to e_2 \\ \rightsquigarrow v, \mathcal{H}' \mid K_1^{\mathrm{matC}} \cdot (q, q') \cdot K_2^{\mathrm{matC}} \end{array}}\ \textsc{E:MatC}$$

Fig. 1. Big-step operational semantics.

that resources are neither used nor restituted. The operation $(q,q')\cdot(p,p')$ defines how to account for an evaluation consisting of evaluations whose resource consumptions are defined by (q,q') and (p,p'), respectively. We define

$$(q,q')\cdot(p,p') = \begin{cases} (q+p-q',\ p') \text{ if } q' \le p \\ (q,\ p'+q'-p) \text{ if } q' > p \end{cases}$$

The intuition is that we need q resource units to perform the first evaluation and after the evaluation q' restituted units remain. Now we have to pay for the second operation which needs p units. If $q' \le p$ then we additionally need $p-q'$ resources to pay for both evaluations and have p' resources left in the end. If $q' > p$ then q units suffices to perform both evaluations. Additionally, the $q'-p$ units that are not needed for the second evaluation are added to the resources becoming finally available.

The following facts are often used in proofs.

Proposition 1. *Let* $(q,q') = (r,r')\cdot(s,s')$.

1. $q \ge r$ *and* $q - q' = r - r' + s - s'$
2. *If* $(p,p') = (\bar{r},r')\cdot(s,s')$ *and* $\bar{r} \ge r$ *then* $p \ge q$ *and* $p' = q'$
3. *If* $(p,p') = (r,r')\cdot(\bar{s},s')$ *and* $\bar{s} \ge s$ *then* $p \ge q$ *and* $p' \le q'$
4. $(r,r')\cdot((s,s')\cdot(t,t')) = ((r,r')\cdot(s,s'))\cdot(t,t')$

If resources are never restituted (as with time) then we can restrict to elements of the form $(q,0)$ and $(q,0)\cdot(p,0)$ is just $(q+p,0)$.

We identify (positive and negative) rational numbers with elements of $\mathcal{R}$ as follows: $q \ge 0$ denotes $(q,0)$ and $q < 0$ denotes $(0,-q)$. This notation avoids case distinctions in the evaluation rules since the constants K that appear in the rules might be negative. Other than in [9], we store pairs on the stack instead of the heap. This leads to a more natural heap-space behavior.

Partial Evaluations A general shortcoming of classic big-step operational semantics is that it does not provide judgments for evaluations that diverge. This is problematic if one intends to prove statements for all computations (divergent and convergent) that do not go wrong.

A straightforward remedy is to use a small-step semantics to describe computations. But in the context of resource analysis, the use of big-step rules seems to be more favorable. First, big-step rules can more directly axiomatize the resource behavior of compiled code on specific machines. Secondly, it allows for shorter and less syntactic proofs.

Another classic approach [10, 11] is to add divergence rules to the operational semantics that are interpreted coinductively. But then one loses the ability to prove statements by induction on the evaluation which is crucial for the proof of the soundness theorem (Thm. 5). It should also be possible to work with a coinductive definition in the style of Cousot or Leroy [10, 11]. However, coinductive semantics leans itself less well to formulating and proving semantic soundness theorems of the form "if the program is well-typed and the operational semantics

$$\frac{}{\mathcal{V},\mathcal{H} \vdash e \rightsquigarrow \mid 0}\ \textsc{P:Zero} \qquad \frac{b \in \{\mathit{True}, \mathit{False}\}}{\mathcal{V},\mathcal{H} \vdash b \rightsquigarrow \mid K^{\mathrm{bool}}}\ \textsc{P:ConstB} \qquad \frac{}{\mathcal{V},\mathcal{H} \vdash () \rightsquigarrow \mid K^{\mathrm{unit}}}\ \textsc{P:ConstU}$$

$$\frac{n \in \mathbb{Z}}{\mathcal{V},\mathcal{H} \vdash n \rightsquigarrow \mid K^{\mathrm{int}}}\ \textsc{P:ConstI} \qquad \frac{x \in \mathrm{dom}(\mathcal{V})}{\mathcal{V},\mathcal{H} \vdash x \rightsquigarrow \mid K^{\mathrm{var}}}\ \textsc{P:Var} \qquad \frac{x_1, x_2 \in \mathrm{dom}(\mathcal{V})}{\mathcal{V},\mathcal{H} \vdash (x_1, x_2) \rightsquigarrow \mid K^{\mathrm{pair}}}\ \textsc{P:Pair}$$

$$\frac{\mathcal{V}(x) = v \qquad [y_f \mapsto v], \mathcal{H} \vdash e_f \rightsquigarrow \mid q}{\mathcal{V},\mathcal{H} \vdash f(x) \rightsquigarrow \mid K_1^{\mathrm{app}} + q}\ \textsc{P:FunApp} \qquad \frac{\mathcal{V},\mathcal{H} \vdash e_1 \rightsquigarrow \mid q}{\mathcal{V},\mathcal{H} \vdash \mathit{let}\ x = e_1\ \mathit{in}\ e_2 \rightsquigarrow \mid K_1^{\mathrm{let}} + q}\ \textsc{P:Let1}$$

$$\frac{\begin{array}{c}\mathcal{V},\mathcal{H} \vdash e_1 \rightsquigarrow v_1, \mathcal{H}_1 \mid (q, q') \\ \mathcal{V}[x \mapsto v_1], \mathcal{H}_1 \vdash e_2 \rightsquigarrow \mid p \qquad K_1^{\mathrm{let}} \cdot (q, q') \cdot K_2^{\mathrm{let}} \cdot (p, 0) = (r, r')\end{array}}{\mathcal{V},\mathcal{H} \vdash \mathit{let}\ x = e_1\ \mathit{in}\ e_2 \rightsquigarrow \mid r}\ \textsc{P:Let2}$$

$$\frac{\mathcal{V}(x) = \mathit{True} \qquad \mathcal{V},\mathcal{H} \vdash e_t \rightsquigarrow \mid q}{\mathcal{V},\mathcal{H} \vdash \mathit{if}\ x\ \mathit{then}\ e_t\ \mathit{else}\ e_f \rightsquigarrow \mid K_1^{\mathrm{conT}} + q}\ \textsc{P:CondT} \qquad \frac{x_1, x_2 \in \mathrm{dom}(\mathcal{V})}{\mathcal{V},\mathcal{H} \vdash x_1\ \mathit{op}\ x_2 \rightsquigarrow \mid K^{\mathrm{op}}}\ \textsc{P:BinOp}$$

$$\frac{\mathcal{V}(x) = \mathit{False} \qquad \mathcal{V},\mathcal{H} \vdash e_f \rightsquigarrow \mid q}{\mathcal{V},\mathcal{H} \vdash \mathit{if}\ x\ \mathit{then}\ e_t\ \mathit{else}\ e_f \rightsquigarrow \mid K_1^{\mathrm{conF}} + q}\ \textsc{P:CondF} \qquad \frac{x_h, x_t \in \mathrm{dom}(\mathcal{V})}{\mathcal{V},\mathcal{H} \vdash \mathit{cons}(x_h, x_t) \rightsquigarrow \mid K^{\mathrm{cons}}}\ \textsc{P:Cons}$$

$$\frac{\mathcal{V}(x) = (v_1, v_2) \qquad \mathcal{V}[x_1 \mapsto v_1, x_2 \mapsto v_2], \mathcal{H} \vdash e \rightsquigarrow \mid q}{\mathcal{V},\mathcal{H} \vdash \mathit{match}\ x\ \mathit{with}\ (x_1, x_2) \to e \rightsquigarrow \mid K_1^{\mathrm{matP}} + q}\ \textsc{P:MatP} \qquad \frac{}{\mathcal{V},\mathcal{H} \vdash \mathit{nil} \rightsquigarrow \mid K^{\mathrm{nil}}}\ \textsc{P:Nil}$$

$$\frac{\mathcal{V}(x) = \textsc{Null} \qquad \mathcal{V},\mathcal{H} \vdash e_1 \rightsquigarrow \mid q}{\mathcal{V},\mathcal{H} \vdash \mathit{match}\ x\ \mathit{with}\ \mid \mathit{nil} \to e_1 \mid \mathit{cons}(x_h, x_t) \to e_2 \rightsquigarrow \mid K_1^{\mathrm{matN}} + q}\ \textsc{P:MatN}$$

$$\frac{\mathcal{V}(x) = l \qquad \mathcal{H}(l) = (v_h, v_t) \qquad \mathcal{V}[x_h \mapsto v_h, x_t \mapsto v_t], \mathcal{H} \vdash e_2 \rightsquigarrow \mid q}{\mathcal{V},\mathcal{H} \vdash \mathit{match}\ x\ \mathit{with}\ \mid \mathit{nil} \to e_1 \mid \mathit{cons}(x_h, x_t) \to e_2 \rightsquigarrow \mid K_1^{\mathrm{matC}} + q}\ \textsc{P:MatC}$$

Fig. 2. Partial big-step operational semantics.

says X then Y holds". For example, in Leroy's Lemmas 17-22 [11] the coinductive definition appears in the conclusion rather than as a premise.

That is why we use a novel approach to the problem here by defining a *big-step semantics for partial evaluations* that directly corresponds to the rules of the big-step semantics in Fig. 1. It defines a statement of the form $\mathcal{V},\mathcal{H} \vdash e \rightsquigarrow \mid q$ for a stack $\mathcal{V}$, a heap $\mathcal{H}$, $q \in \mathbb{Q}^+$ and an expression e. The meaning is that there is a partial evaluation of e with the stack $\mathcal{V}$ and the heap $\mathcal{H}$ that consumes q resources. Here, q is the watermark of the resource usage. We do not have to keep track of the restituted resources since partial evaluations are composed of complete evaluations only.

Note that the rule P:Zero is essential for the partiality of the semantics. It can be applied at any point to stop the evaluation and thus yields to a non-deterministic evaluation judgment.

Since there might be negative constants K, the partial evaluation rules have conclusions of the form $\mathcal{V},\mathcal{H} \vdash e \rightsquigarrow \mid \max(q,0)$ to ensure non-negative values. For simplicity we just write $\mathcal{V},\mathcal{H} \vdash e \rightsquigarrow \mid q$ instead of $\mathcal{V},\mathcal{H} \vdash e \rightsquigarrow \mid \max(q,0)$ in each conclusion of the rules in Fig. 2.

We prove that if an expression converges in a given environment then the resource-usage watermark of the evaluation is an upper bound for the resource usage of every partial evaluation of the expression in that environment.

Theorem 1. *If* $\mathcal{V},\mathcal{H} \vdash e \rightsquigarrow v,\mathcal{H}' \mid (q,q')$ *and* $\mathcal{V},\mathcal{H} \vdash e \rightsquigarrow \mid p$ *then* $p \leq q$.

Proof. By induction on the derivation D of $\mathcal{V},\mathcal{H} \vdash e \rightsquigarrow v,\mathcal{H}' \mid (q,q')$. To prove the *induction basis* let D consist of one step. If e was derived by P:Zero then $0 = p \leq q$ since $q \geq 0$. Otherwise e is a constant c, a variable x, a binary operation x_1 *op* x_2, a pair (x_1,x_2), the constant *nil*, or *cons*(x_1,x_2). For example, let e be a variable x. Then by definition of E:Var $\mathcal{V},\mathcal{H} \vdash e \rightsquigarrow v,\mathcal{H}' \mid (K^{\text{var}},0)$. But the only P-rules that apply to x are P:Var and P:Zero. Thus it follows that if $\mathcal{V},\mathcal{H} \vdash e \rightsquigarrow \mid p$ then $p = 0$ or $p = K^{\text{var}}$. The other cases are similar.

For the induction step assume that $|D| > 1$. Then e is a pattern match, a function application, a conditional, or a let expression. For instance, let e be the expression *let* $x = e_1$ *in* e_2. Then it follows from rule E:Let that $\mathcal{V},\mathcal{H} \vdash e_1 \rightsquigarrow v_1,\mathcal{H}_1 \mid (q_1,q_1')$, $\mathcal{V}[x \mapsto v_1],\mathcal{H}_1 \vdash e_2 \rightsquigarrow v_2,\mathcal{H}_2 \mid (q_2,q_2')$ and

$$(q,q') = K_1^{\text{let}} \cdot (q_1,q_1') \cdot K_2^{\text{let}} \cdot (q_2,q_2') \cdot K_3^{\text{let}} \tag{1}$$

By induction we conclude

$$\text{if } \mathcal{V},\mathcal{H} \vdash e_1 \rightsquigarrow \mid p_1 \text{ then } p_1 \leq q_1 \tag{2}$$

$$\text{if } \mathcal{V}[x \mapsto v_1],\mathcal{H}_1 \vdash e_2 \rightsquigarrow \mid p_2 \text{ then } p_2 \leq q_2 \tag{3}$$

Now let $\mathcal{V},\mathcal{H} \vdash e \rightsquigarrow \mid p$. Then this statement was derived via the rules P:Let1 or P:Let2. In the first case it follows from (2) and (1) that $p \leq q_1 + K_1^{\text{let}} \leq q$.

If $\mathcal{V},\mathcal{H} \vdash e \rightsquigarrow \mid p$ was derived by P:Let2 then it follows that $(p,p') = K_1^{\text{let}} \cdot (q_1,q_1') \cdot K_2^{\text{let}} \cdot (p_2,0)$ for some p',p_2. We conclude from (3) that $p_2 \leq q_2$ and hence from Prop. 1 and (1) $p \leq q$. The other cases are similar to the case P:Let1. □

A stack $\mathcal{V}$ and a heap $\mathcal{H}$ are *well-formed* with respect to a context Γ if, for every $x \in \text{dom}(\Gamma)$, $\mathcal{V}(x)$ is a value matching the type $\Gamma(x)$ or a location in $\mathcal{H}$ that contains a value matching $\Gamma(x)$. We then write $\mathcal{H} \vDash \mathcal{V}{:}\Gamma$. Similarly, we write $\mathcal{H} \vDash v{:}A$ if v is a value matching type A in $\mathcal{H}$. A formal definition is given in [7]. Thm. 2 shows that the evaluation of a well-typed expression in a well-formed environment results in a well-formed environment.

Theorem 2. *If* $\Gamma \vdash_\Sigma e{:}A$, $\mathcal{H} \vDash \mathcal{V}{:}\Gamma$ *and* $\mathcal{V},\mathcal{H} \vdash e \rightsquigarrow v,\mathcal{H}' \mid (q,q')$ *then* $\mathcal{H}' \vDash \mathcal{V}{:}\Gamma$ *and* $\mathcal{H}' \vDash v{:}A$.

Proof. Induction on the derivation of $\mathcal{V},\mathcal{H} \vdash e \rightsquigarrow v,\mathcal{H}' \mid (q,q')$. □

Thm. 3 states that, in a well-formed environment, every well-typed expression either diverges or evaluates to a value of the stated type. To this end we instantiate the resource constants in the rules to count the number of evaluation steps. We first prove a lemma that shows that there is a partial evaluation that uses the same number of steps as the complete evaluation. It is used in the proof of Thm. 3.

Lemma 1. *Let the resource constants be instantiated by* $K^x = 1$*,* $K^x_1 = 1$ *and* $K^x_m = 0$ *for all* x *and all* $m > 1$*. If* $\mathcal{V}, \mathcal{H} \vdash e \rightsquigarrow v, \mathcal{H}' \mid (n, 0)$ *then* $\mathcal{V}, \mathcal{H} \vdash e \rightsquigarrow \mid n$*.*

Proof. By induction on the derivation of $\mathcal{V}, \mathcal{H} \vdash e \rightsquigarrow v, \mathcal{H}' \mid (n, 0)$. It is essential that $K^x_m = 0$ for all x and all $m > 1$. □

Theorem 3. *Let the resource constants be instantiated by* $K^x = 1$*,* $K^x_1 = 1$ *and* $K^x_m = 0$ *for all* x *and all* $m > 1$*. If* $\Gamma \vdash_\Sigma e{:}A$ *and* $\mathcal{H} \vDash \mathcal{V}{:}\Gamma$ *then* $\mathcal{V}, \mathcal{H} \vdash e \rightsquigarrow v, \mathcal{H}' \mid (n, 0)$ *for an* $n \in \mathbb{N}$ *or* $\mathcal{V}, \mathcal{H} \vdash e \rightsquigarrow \mid m$ *for every* $m \in \mathbb{N}$*.*

Proof. We show by induction on n that if

$$\Gamma \vdash_\Sigma e{:}A,\ \mathcal{V}, \mathcal{H} \vdash e \rightsquigarrow \mid n \text{ and } \mathcal{H} \vDash \mathcal{V}{:}\Gamma \tag{4}$$

then $\mathcal{V}, \mathcal{H} \vdash e \rightsquigarrow v, \mathcal{H}' \mid (n, 0)$ or $\mathcal{V}, \mathcal{H} \vdash e \rightsquigarrow \mid n + 1$. Then Thm. 3 follows since $\mathcal{V}, \mathcal{H} \vdash e \rightsquigarrow \mid 0$ for every $\mathcal{V}, \mathcal{H}$ and e.

Induction basis $n = 0$: We conclude from the well-formedness of the environment (4) that $\mathcal{V}, \mathcal{H} \vdash e \rightsquigarrow \mid 1$ (by case distinction over the partial evaluation rules).

Induction step $n > 0$: Assume (4). If e is a constant c, a variable x, a binary operation x_1 *op* x_2, a pair (x_1, x_2), the constant *nil*, or *cons*(x_1, x_2). Then $n = 1$ and we conclude $\mathcal{V}, \mathcal{H} \vdash e \rightsquigarrow v, \mathcal{H}' \mid (1, 0)$ directly from the corresponding evaluation rule. If e is a pattern match, a function application, a conditional, or a let expression then we use the induction hypothesis.

Since the other cases are similar, we provide the argument only for the case where e is a let expression *let* $x = e_1$ *in* e_2. Then $\mathcal{V}, \mathcal{H} \vdash e \rightsquigarrow \mid n$ was derived via P:Let1 or P:Let2. In the case of P:Let1 it follows that $\mathcal{V}, \mathcal{H} \vdash e_1 \rightsquigarrow \mid n - 1$. By the induction hypothesis we conclude that either $\mathcal{V}, \mathcal{H} \vdash e_1 \rightsquigarrow \mid n$ or $\mathcal{V}, \mathcal{H} \vdash e_1 \rightsquigarrow v_1, \mathcal{H}_1 \mid (n - 1, 0)$. In the first case we can use P:Let1 to derive $\mathcal{V}, \mathcal{H} \vdash e \rightsquigarrow \mid n + 1$. In the second case it follows from Thm. 2 that $\mathcal{H}_1 \vDash \mathcal{V}{:}\Gamma$ and $\mathcal{H}_1 \vDash v_1{:}A$ and thus $\mathcal{H}_1 \vDash \mathcal{V}[x \mapsto v_1]{:}\Gamma, x{:}A$. Like in the induction basis we have then $\mathcal{V}[x \mapsto v_1], \mathcal{H}_1 \vdash e_2 \rightsquigarrow \mid 1$. Therefore we can apply P:Let2 to obtain $\mathcal{V}, \mathcal{H} \vdash e \rightsquigarrow \mid n + 1$.

Assume now that e was derived by the use of P:Let2. Then it is true that $\mathcal{V}, \mathcal{H} \vdash e_1 \rightsquigarrow v_1, \mathcal{H}_1 \mid (n_1, 0)$ and $\mathcal{V}[x \mapsto v_1], \mathcal{H}_1 \vdash e_2 \rightsquigarrow \mid n_2$ for some n_1, n_2 with $n_1 + n_2 + 1 = n$. From Thm. 2 it follows that $\mathcal{H}_1 \vDash \mathcal{V}[x \mapsto v_1]{:}\Gamma, x{:}A$. Therefore we can apply the induction hypothesis to infer that $\mathcal{V}[x \mapsto v_1], \mathcal{H}_1 \vdash e_2 \rightsquigarrow v_2, \mathcal{H}_2 \mid (n_2, 0)$ or $\mathcal{V}[x \mapsto v_1], \mathcal{H}_1 \vdash e_2 \rightsquigarrow \mid n_2 + 1$. In the first case we apply E:Let and obtain $\mathcal{V}, \mathcal{H} \vdash e \rightsquigarrow v_2, \mathcal{H}_2 \mid (n, 0)$. In the second case we apply P:Let2 and obtain $\mathcal{V}, \mathcal{H} \vdash e \rightsquigarrow \mid n + 1$. □

Cost-Free Metric The type inference algorithm makes use of the *cost-free* resource metric. This is the metric in which all constants K that appear in the rules are instantiated to zero. We will use it in §5 to define a resource-polymorphic recursion where we use cost-free function types to pass potential from the argument to the result. The following proposition is direct.

Proposition 2. *Let all resource constants K be instantiated by $K = 0$. If $\mathcal{V}, \mathcal{H} \vdash e \rightsquigarrow v, \mathcal{H}' \mid (q, q')$ then $q = q' = 0$. If $\mathcal{V}, \mathcal{H} \vdash e \rightsquigarrow \mid q$ then $q = 0$.*

5 Resource Annotated Types

Resource-annotated types are simple types where lists are annotated with non-negative vectors $\vec{p} \in \mathbb{Q}^n$. Here we only give a short definition of the potential functions defined by annotated types. More explanations can be found in [9].

Let $\vec{p} = (p_1, \ldots, p_k)$ be an annotation for a list type. The *additive shift* of $\vec{p}$ is $\lhd(\vec{p}) = (p_1 + p_2, p_2 + p_3, \ldots, p_{k-1} + p_k, p_k)$. Let $\mathcal{H}$ be a heap, A be a resource-annotated type and let v be a value matching type A in $\mathcal{H}$. The *potential* $\Phi_{\mathcal{H}}(v{:}A)$ is then defined as follows.

$$\Phi_{\mathcal{H}}(v{:}A) = 0 \text{ if } v = \textsc{Null} \text{ or } A \in \{\mathit{unit}, \mathit{int}, \mathit{bool}\}$$
$$\Phi_{\mathcal{H}}((v_1, v_2){:}(A_1, A_2)) = \Phi_{\mathcal{H}}(v_1{:}A_1) + \Phi_{\mathcal{H}}(v_2{:}A_2)$$
$$\Phi_{\mathcal{H}}(l{:}L^{\vec{p}}(A')) = p_1 + \Phi_{\mathcal{H}}(v'{:}A') + \Phi_{\mathcal{H}}(l'{:}\ L^{\lhd(\vec{p})}(A')) \text{ if } \mathcal{H}(l){=}(v', l')$$

If l_1 is a location that points to a list then we write $\mathcal{H}(l_1) = [v_1, \ldots, v_n]$ if $\mathcal{H}(l_i) = (v_i, l_{i+1})$ for $i = 1, \ldots, n$ and $l_{n+1} = \textsc{Null}$. If $l_1 = \textsc{Null}$ then $\mathcal{H}(l_1) = []$. Thm. 4 shows how to express the potential $\Phi_{\mathcal{H}}(v{:}A)$ of a value v with respect the heap $\mathcal{H}$ and a matching annotated type A in terms of polynomials in the lengths of the lists that are reachable from v. A proof can be found in the extended version of [9].

Theorem 4. *Let $\mathcal{H}$ be a heap and let $\mathcal{H}(l) = [v_1 \ldots, v_n]$ be a list of length n. Then $\Phi_{\mathcal{H}}(l{:}L^{\vec{p}}(A)) = \sum_{i=1}^{k} p_i \binom{n}{i} + \sum_{i=1}^{n} \Phi_{\mathcal{H}}(v_i{:}A)$.*

As in the case of simple types, a *typing context* is a finite partial mapping from variable identifiers to annotated data types. The potential of a context Γ with respect to a heap $\mathcal{H}$ and a stack $\mathcal{V}$ is $\Phi_{\mathcal{V},\mathcal{H}}(\Gamma) = \sum_{x \in \mathrm{dom}(\Gamma)} \Phi_{\mathcal{H}}(\mathcal{V}(x){:}\Gamma(x))$.

Resource-annotated first-order types have the form $A \xrightarrow{q/q'} B$ for $q, q' \in \mathbb{Q}^+$ and annotated data types A, B. A *resource-annotated signature* Σ is a finite, partial mapping from function identifiers to resource-annotated first-order types. A *resource-annotated typing judgment* has the form $\Sigma; \Gamma \; {}^{k}\!\vdash^{q}_{q'} e{:}A$ where e is a RAML expression, $k \in \mathbb{N}^+$ is the length of the list annotations, $q, q' \in \mathbb{Q}^+$ are non-negative rational numbers, Σ is a resource-annotated signature, Γ is a resource-annotated context and A is a resource-annotated data type. The intended meaning of this judgment is that if there are more than $q + \Phi(\Gamma)$ resource units available then this is sufficient to evaluate e and there are more than $q' + \Phi(v{:}A)$ resource units if e evaluates to the value v.

$$\frac{n \in \mathbb{Z} \qquad q \geq q' + K^{\mathrm{int}}}{\varGamma \ {}^{k}\!\vdash^{q}_{q'}\ n : \mathit{int}}\ \textsc{T:ConstI} \qquad \frac{b \in \{\mathit{True}, \mathit{False}\} \qquad q \geq q' + K^{\mathrm{bool}}}{\varGamma \ {}^{k}\!\vdash^{q}_{q'}\ b{:}\mathit{bool}}\ \textsc{T:ConstB}$$

$$\frac{q \geq q' + K^{\mathrm{unit}}}{\varGamma \ {}^{k}\!\vdash^{q}_{q'}\ (){:}\mathit{unit}}\ \textsc{T:ConstU} \qquad \frac{\mathit{op} \in \{+,-,*,\mathit{mod},\mathit{div}\} \qquad q \geq q' + K^{\mathrm{op}}}{\varGamma, x_1{:}\mathit{int}, x_2{:}\mathit{int} \ {}^{k}\!\vdash^{q}_{q'}\ x_1\ \mathit{op}\ x_2 : \mathit{int}}\ \textsc{T:BinOpI}$$

$$\frac{\mathit{op} \in \{\mathit{or}, \mathit{and}\} \qquad q \geq q' + K^{\mathrm{op}}}{\varGamma, x_1{:}\mathit{bool}, x_2{:}\mathit{bool} \ {}^{k}\!\vdash^{q}_{q'}\ x_1\ \mathit{op}\ x_2 : \mathit{bool}}\ \textsc{T:BinOpB} \qquad \frac{q \geq q' + K^{\mathrm{var}}}{\varGamma, x{:}A \ {}^{k}\!\vdash^{q}_{q'}\ x : A}\ \textsc{T:Var}$$

$$\frac{k = 1 \qquad \varSigma(f) = B \xrightarrow{p/p'} A \qquad q = p{+}c{+}K^{\mathrm{app}}_1 \qquad q' = p'{+}c{-}K^{\mathrm{app}}_2}{\varGamma, x{:}B \ {}^{k}\!\vdash^{q}_{q'}\ f(x) : A}\ \textsc{T:FunApp1}$$

$$\frac{\begin{array}{c} q = p + p_{cf} + c + K^{\mathrm{app}}_1 \qquad q' = p' + p'_{cf} + c - K^{\mathrm{app}}_2 \\ \varSigma_{cf}; y_f{:}B_{cf} \ {}^{cf(k-1)}\!\vdash^{p_{cf}}_{p'_{cf}}\ e_f{:}A_{cf} \qquad \varSigma_{cf}(f) = B_{cf} \xrightarrow{p_{cf}/p'_{cf}} A_{cf} \\ \curlyvee(A \mid A', A_{cf}) \qquad \curlyvee(B \mid B', B_{cf}) \qquad \varSigma(f) = B' \xrightarrow{p/p'} A' \end{array}}{\varGamma, x{:}B \ {}^{k}\!\vdash^{q}_{q'}\ f(x) : A}\ \textsc{T:FunApp}$$

$$\frac{\begin{array}{c} q \geq p_1{+}K^{\mathrm{let}}_1 \qquad p'_1 \geq p_2{+}K^{\mathrm{let}}_2 \qquad p'_2 \geq q'{+}K^{\mathrm{let}}_3 \qquad \curlyvee(\varDelta \mid \varDelta_1, \varDelta_2) \\ \mathrm{Var}(\varGamma_1) \cap \mathrm{Var}(\varGamma_2) = \emptyset \qquad \varGamma_1, \varDelta_1 \ {}^{k}\!\vdash^{p_1}_{p'_1}\ e_1{:}B \qquad \varGamma_2, \varDelta_2, x{:}B \ {}^{k}\!\vdash^{p_2}_{p'_2}\ e_2{:}A \end{array}}{\varGamma_1, \varGamma_2, \varDelta \ {}^{k}\!\vdash^{q}_{q'}\ \mathit{let}\ x = e_1\ \mathit{in}\ e_2 : A}\ \textsc{T:Let}$$

$$\frac{\begin{array}{c} q \geq p_t{+}K^{\mathrm{conT}}_1 \qquad q \geq p_f{+}K^{\mathrm{conF}}_1 \qquad p'_t \geq q'{+}K^{\mathrm{conT}}_2 \qquad p'_f \geq q'{+}K^{\mathrm{conF}}_2 \\ A_i <: A \text{ for } i = 1,2 \qquad \varGamma \ {}^{k}\!\vdash^{p_t}_{p'_t}\ e_t : A_1 \qquad \varGamma \ {}^{k}\!\vdash^{p_f}_{p'_f}\ e_f : A_2 \end{array}}{\varGamma, x{:}\mathit{bool} \ {}^{k}\!\vdash^{q}_{q'}\ \mathit{if}\ x\ \mathit{then}\ e_t\ \mathit{else}\ e_f : A}\ \textsc{T:Cond}$$

$$\frac{q \geq p + K^{\mathrm{matP}}_1 \qquad p' \geq q' + K^{\mathrm{matP}}_2 \qquad \varGamma, x_1{:}B_1, x_2{:}B_2 \ {}^{k}\!\vdash^{p}_{p'}\ e{:}A}{\varGamma, x{:}(B_1, B_2) \ {}^{k}\!\vdash^{q}_{q'}\ \mathit{match}\ x\ \mathit{with}\ (x_1, x_2) \rightarrow e : A}\ \textsc{T:MatP}$$

$$\frac{q \geq q' + K^{\mathrm{pair}}}{\varGamma, x_1{:}A_1, x_2{:}A_2 \ {}^{k}\!\vdash^{q}_{q'}\ (x_1, x_2) : (A_1, A_2)}\ \textsc{T:Pair} \qquad \frac{q \geq q' + K^{\mathrm{nil}}}{\varGamma \ {}^{k}\!\vdash^{q}_{q'}\ \mathit{nil}{:}L(A)}\ \textsc{T:Nil}$$

$$\frac{\vec{p} = (p_1, \ldots, p_k) \qquad \vec{r} \geq \lhd(\vec{p}) \qquad q \geq q'{+}p_1{+}K^{\mathrm{cons}} \qquad A_i <: A \text{ for } i = 1,2}{\varGamma, x_h{:}A_1, x_t{:}L^{\vec{r}}(A_2) \ {}^{k}\!\vdash^{q}_{q'}\ \mathit{cons}(x_h, x_t){:}L^{\vec{p}}(A)}\ \textsc{T:Cons}$$

$$\frac{\begin{array}{c} \vec{p} = (p_1, \ldots, p_k) \qquad A_i <: A \text{ for } i = 1,2 \\ q + p_1 \geq s_c + K^{\mathrm{matC}}_1 \qquad q \geq s_n + K^{\mathrm{matN}}_1 \qquad s'_c \geq q' + K^{\mathrm{matC}}_2 \\ s'_n \geq q' + K^{\mathrm{matN}}_2 \qquad \varGamma \ {}^{k}\!\vdash^{s_n}_{s'_n}\ e_1{:}A_1 \qquad \varGamma, x_h{:}B, x_t{:}L^{\lhd(\vec{p})}(B) \ {}^{k}\!\vdash^{s_c}_{s'_c}\ e_2{:}A_2 \end{array}}{\varGamma, x{:}L^{\vec{p}}(B) \ {}^{k}\!\vdash^{q}_{q'}\ \mathit{match}\ x\ \mathit{with}\ \mid \mathit{nil} \rightarrow e_1 \mid \mathit{cons}(x_h, x_t) \rightarrow e_2 : A}\ \textsc{T:MatL}$$

Fig. 3. Algorithmic type rules.

A RAML program with resource-annotated types of degree k consists of a resource-annotated signature Σ and a family of expressions with variables identifiers $(e_f, y_f)_{f \in \mathrm{dom}(\Sigma)}$ such that for each e_f we have $\Sigma; y_f{:}A \ {}^{k}\!\vdash^{q}_{q'} e_f{:}B$ if $\Sigma(f){=}A \xrightarrow{q/q'} B$.

We write $\Sigma; \Gamma \ {}^{cf(k)}\!\vdash^{q}_{q'} e{:}A$ to refer to cost-free type judgments where all constants K in the rules are zero. As described in §2 we use it to define a resource-polymorphic recursion where we use cost-free function types to pass potential from the argument to the result.

In the typing rules in Fig. 3 we write $e[z/x]$ to denote the expression e with all free occurrences of the variable x replaced with the variable z. We assume that a fixed but arbitrary global resource-annotated signature Σ is given. Furthermore, there is the implicit constraint $q \geq 0$ for every resource annotation q.

The rules are mostly algorithmic versions of the typing rules in [9]. The most important difference is the rule T:FunApp which enables resource-polymorphic recursion. It states that one can add any cost-free typing of the function body to the function type that is given by the signature Σ. Note that $(e_f, y_f)_{f \in \Sigma_{cf}}$ must be a valid RAML program with cost-free types of degree $k-1$. The annotated signature Σ_{cf} used can differ in every application of the rule.

The idea is as follows. In order to pay for the resource costs of a function call $f(x)$, the available potential $(\Phi(x{:}B) + q)$ must meet the requirements of the functions' signature $(\Phi(x{:}B') + p)$. Additionally available potential $(\Phi(x{:}B_{cf}) + p_{cf})$ can be passed to a cost-free typing of the function body. The potential after the function call $(\Phi(f(x){:}A) + q')$ is then the sum of the potentials that are assigned by the cost-free typing $(\Phi(f(x){:}A_{cf}){+}p_{cf})$ and by the function signature $(\Phi(f(x){:}A') + p)$. As a result, $f(x)$ can be used resource-polymorphically with a specific typing for each recursive call while the resource monomorphic function signature enables an efficient type inference.

The *sharing relation* $\curlyvee$ defines how potential can be shared between multiple occurrences of a variable. We have $\curlyvee(A \mid A_1, A_2)$ only if A, A_1 and A_2 are structurally identical and $\Phi(v{:}A) = \Phi(v{:}A_1) + \Phi(v{:}A_2)$ for every value v.

$$\curlyvee(C \mid C, C) \text{ if } C \in \{\mathit{unit}, \mathit{bool}, \mathit{int}\}$$
$$\curlyvee(L^{\vec{p}}(A) \mid L^{\vec{q}}(A_1), L^{\vec{r}}(A_2)) \text{ if } \curlyvee(A \mid A_1, A_2) \text{ and } \vec{p} = \vec{q} + \vec{r}$$
$$\curlyvee((A, B) \mid (A_1{,}B_1), (A_2, B_2)) \text{ if } \curlyvee(X \mid X_1, X_2) \text{ for } X{=}A, B$$

We define $\curlyvee(\Gamma \mid \Gamma_1, \Gamma_2)$ iff $\Gamma = x_1{:}A_1, \ldots, x_n{:}A_n$, $\Gamma_i = x_1{:}A^i_1, \ldots, x_n{:}A^i_n$ and $\curlyvee(A_i \mid A^1_i, A^2_i)$ for all $i \in \{1, \ldots, n\}$.

A data type A is a *subtype* of a data type B only if A and B are structurally identical, and for every value v the potential of $v{:}A$ is greater or equal than the potential of $v{:}B$. We then write $A <: B$. Formally, we define

$$C <: C \text{ if } C \in \{\mathit{unit}, \mathit{bool}, \mathit{int}\}$$
$$(A_1, A_2) <: (B_1, B_2) \text{ if } A_1 <: B_1 \text{ and } A_2 <: B_2$$
$$L^{\vec{p}}(A) <: L^{\vec{q}}(B) \text{ if } A <: B \text{ and } \vec{p} \geq \vec{q}$$

The introduction of the partial evaluation rules enables us to formulate a stronger soundness theorem than, e.g., in [9]. It states that the bounds derived from an annotated type statement also hold for non-terminating evaluations. Additionally, the new notation that we use in the operational semantics allows for a more concise statement.

Theorem 5 (Soundness). *Let $\mathcal{H} \vDash \mathcal{V}{:}\Gamma$ and let $\Gamma \;{}^{k}\!\vdash^{q}_{q'}\; e{:}A$.*

1. *If $\mathcal{V},\mathcal{H} \vdash e \rightsquigarrow v,\mathcal{H}' \mid (p,p')$ then $p \leq \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + q$ and $p - p' \leq \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + q - (\Phi_{\mathcal{H}'}(v{:}A) + q')$.*
2. *If $\mathcal{V},\mathcal{H} \vdash e \rightsquigarrow \mid p$ then $p \leq \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + q$.*

It follows from Thm. 5 and Thm. 3 that run-time bounds also prove the termination of programs.

Corollary 1. *Let the resource constants be instantiated by $K^x = 1$, $K^x_1 = 1$ and $K^x_m = 0$ for all x and all $m > 1$. If $\mathcal{H} \vDash \mathcal{V}{:}\Gamma$ and $\Gamma \;{}^{k}\!\vdash^{q}_{q'}\; e{:}A$ then there is an $n \in \mathbb{N}, n \leq \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + q$ such that $\mathcal{V},\mathcal{H} \vdash e \rightsquigarrow v,\mathcal{H}' \mid (n,0)$.*

Thm. 5 is proved by induction on the derivation of the evaluation statements $\mathcal{V},\mathcal{H} \vdash e \rightsquigarrow v,\mathcal{H}' \mid (p,p')$ and $\mathcal{V},\mathcal{H} \vdash e \rightsquigarrow \mid p$, respectively. There is one proof for all possible instantiations of the resource constants. It is technically involved but conceptually unsurprising. Compared to earlier works [7, 9], further complexity arises from the matching of the constraints in the type rules with the monoid elements in the semantics.

The proof can be found in the appendix. The following two lemmas are needed. They follow directly from the definitions.

Lemma 2. *Let $\mathcal{H} \vDash \mathcal{V}{:}\Gamma$ and let $\mathcal{H} \vDash v{:}A$. If $\curlyvee(A \mid A_1, A_2)$ then $\Phi_{\mathcal{V},\mathcal{H}}(v{:}A) = \Phi_{\mathcal{V},\mathcal{H}}(v{:}A_1) + \Phi_{\mathcal{V},\mathcal{H}}(v{:}A_2)$.*

Lemma 3. *Let $\mathcal{H} \vDash \mathcal{V}{:}\Gamma$ and let $\mathcal{H} \vDash v{:}A$. If $A <: B$ then $\Phi_{\mathcal{V},\mathcal{H}}(v{:}A) \geq \Phi_{\mathcal{V},\mathcal{H}}(v{:}B)$.*

Lemma 4 is used to show the soundness of the rule T:Let. It is proved by induction on the evaluation of the expression e. With the language features that we describe in this paper, the potential of a context is invariant during the evaluation since allocated heap-cells are immutable. So we could replace the last statement of the lemma with $\Phi_{\mathcal{V},\mathcal{H}}(\Gamma) = \Phi_{\mathcal{V},\mathcal{H}'}(\Gamma)$. However, we state the weaker variant of the lemma which remains true in the present of a destructive pattern match. Intuitively, the deletion (deallocation) of heap cells can lead to a reduction of potential.

Lemma 4. *Let $\mathcal{H} \vDash \mathcal{V}{:}\Gamma$, $\Gamma \;{}^{k}\!\vdash^{q}_{q'}\; e{:}A$ and $\mathcal{V},\mathcal{H} \vdash e \rightsquigarrow v,\mathcal{H}' \mid (p,p')$. Then $\Phi_{\mathcal{V},\mathcal{H}}(\Gamma) \geq \Phi_{\mathcal{V},\mathcal{H}'}(\Gamma)$.*

6 The Inference Algorithm

The inference algorithm is mainly defined by the type rules in the previous section. It works like a standard type inference in which each type is annotated with resource variables and the corresponding linear constraints are collected as each type rule is applied. The main innovation in comparison to the classic algorithm [1] is the resource-polymorphic recursion enabled by the rule T:FunApp.

The number of computed constraints grows linearly in the maximal degree k that has to be provided by the user. There is a trade-off between the quality of the analysis and the size of the constraint system. The reason is that one sometimes has to analyze function applications context-sensitively with respect to the call stack. Remember for instance the expression *attach(x,attach(y,xs))* from §1 where we had to use two different types for *attach*.

In our implementation we collapse the cycles in the call graph and analyze each function once for every path in the resulting graph. In a nutshell, the algorithm computes inequalities for annotations of degree k for a strongly connected component (SCC) F of the call graph as follows.

1. Annotate the signature of each function $f \in F$ with fresh resource variables.
2. Use the type rules from §5 to type the corresponding expressions e_f. Introduce fresh resource variables for each type annotation in the derivation and collect the corresponding inequalities.
   (a) For a function application $g \in F$: if $k = 1$ or in the cost-free case use the function resource-monomorphically with the signature from (1). Otherwise, go to (1) and derive a cost-free typing of e_g with a fresh signature. Store the arising inequalities and use the resource variables from the obtained typing together with the signature from (1) in T:FunApp.
   (b) For a function application $g \notin F$: repeat the algorithm for the SSC of g. Store the arising inequalities and use the obtained annotated type of g.

The context sensitivity can lead to an exponential blow up of the constraint system if there is a sequence of function $f_1, \ldots, f_n$ such that f_i calls f_{i+1} several times. But such sequences are very short in most programs. It would not be a substantial limitation in practice to restrict oneself to programs that feature a collapsed call graph with a fixed maximal path length to certainly obtain a constraint system that is linear in the program size.

In general, the computed constraint systems are simple and can be quickly solved by standard LP-solvers. The objective function states that annotations of arguments in function signatures have to be minimized and that annotations of high degree are more expensive then annotations of low degree.

Table 1 shows some computed bounds together with the run time of the prototype implementation on a MacBook Pro with a 2.16 GHz Intel Core 2 Duo. The analysed code is available online. The inference algorithm works very efficiently and infers resource-polymorphic types for all programs that we manually typed in our system. However, it is not complete with respect to full resource-polymorphism. This would mean to start with a (possibly infinite) set of annotated function types for each function and to justify each type with a type

	Computed Evaluation-Step Bound	Act. Behavior	Run Time
quicksort*	$3 + 14a + 12a^2$	$O(a^2)$	0.1438 s
insertionsort*	$3 + 6a + 6a^2$	$O(a^2)$	0.0542 s
mergesort	$5 - 46a + 46a^2$	$O(a \log a)$	0.3059 s
pairs*	$3 + 7a + 9a^2$	$O(a^2)$	0.0507 s
triples*	$3 + 24a - 10a^2 + 6a^3$	$O(a^3)$	0.3043 s
quadruples*	$3 + 1.8\overline{3}a + 26.75a^2 - 10.8\overline{3}a^3 + 2.25a^4$	$O(a^4)$	0.5527s
dyade	$3 - a + 11a^2 + 29b + 5b^2$	$O(a{\cdot}b)$	0.1673 s
lcs	$13 - 22.5a + 24.5a^2 + 67.5b + 30.5b^2$	$O(a{\cdot}b)$	0.8460 s
eratos*	$3 + 4a + 8a^2$	$O(a^2)$	0.0447 s
startBreadth	$17 + 45a + 45a^2$	$O(a^2)$	0.8113 s

Table 1. The computed evaluation-step bounds, the actual worst-case time behavior, and the run time of the analysis in seconds. Experiments with sample inputs of various sizes showed that the bounds are tight for the functions marked with a star.

derivation that uses some first-order types from the initial set. Such a derivation exists for the function *round* below.

round l = **match** l **with** | nil → nil | (x::xs) → x::double (round (half xs))

The function *half* deletes every second element and *double* doubles every element a list. On can assign the cost-free types *half:* $L^1(\mathit{unit})\xrightarrow{0/0}L^2(\mathit{unit})$ and *double:* $L^2(\mathit{unit})\xrightarrow{0/0}L^1(\mathit{unit})$. To derive *round:* $L^a(\mathit{unit})\xrightarrow{0/0}L^a(\mathit{unit})$ for $a = 1$ one would need the type resource-polymorphic linear type with $a = 2$. Since the linear cost-free type already requires resource polymorphism, our algorithm can not infer a typing for *round*. For every $q \in \mathbb{Q}^+$ one can create functions where one would need to multiply some resource annotations with q in cost-free typing of the recursive call. So it is unlikely that there is a method to infer a typing for such functions with a method that uses only linear constraints. One could move to quadratic constraints to address the problem but the efficiency of such an approach is unclear. We plan to also experiment with SMT solvers to deal which such constraints.

7 Case Study: Sorting Algorithms in RAML

A classic application of quantitative resource analysis is the run-time analysis of sorting algorithms. In the book *The Art of Computer Programming* [12], Knuth manually determines worst-case bounds for many well-known sorting algorithms that are implemented in an assembly language for the MIX architecture. Among the analyzed algorithms are quick sort, which uses at most $2n^2 + 37n + 3$ MIX cycles, insertion sort, at most $9\binom{n}{2} + 7n - 6 = 4.5n^2 + 2.5n - 6$ MIX cycles, and merge sort, roughly $10n \log n + 4.92n$ MIX cycles[3] (n is the size of the input

[3] The actual worst-case bound is more complicated and presented in a form that is only meaningful in combination with the source code.

data). As a result of a careful and elaborate analysis, the bounds are tight in the sense that they exactly match the actual worst-case behavior of the functions.

In this section we implement the three sorting algorithms in RAML to automatically determine a bound on the number of evaluation steps they use. We then compare our automatic analysis with the manual analysis of Knuth.

Insertion Sort Below is the implementation of insertion sort in RAML. The same implementation may also be given in a textbook.

```
insert (x,l)  = match l with | nil → [x]
              | (y::ys) → if y < x then y::insert(x,ys) else x::y::ys;

isort l = match l with | nil → nil | (x::xs) → insert (x,isort xs);
```

If we instantiate our type system with the evaluation-step metric then the prototype implementation automatically computes the following types.

insert: $(int, L^{(12,0)}(int)) \xrightarrow{5/0} L^{(0,0)}(int)$ isort: $L^{(12,12)}(int) \xrightarrow{3/0} L^{(0,0)}(int)$

This means that *insert* needs at most $5 + 12n$ evaluation steps and *isort* needs at most $3 + 6n + 6n^2$ if n is the size of the respective input list.[4] In the type derivation of *isort* we need resource-polymorphic recursion since the result of the recursive call has to contain potential to pay for the following evaluation of *insert*. The type of the recursive call is *isort:*$L^{(24,12)}(int) \xrightarrow{3/0} L^{(12,0)}(int)$.

Quick Sort Quick sort can also be implemented in RAML in the usual way.

```
append(l,ys) = match l with | nil → ys | (x::xs) → x::append(xs,ys);

split (p,l)  = match l with | nil → (nil,nil)
             | (x::xs) → let (ls,rs) = split (p,xs) in
                         if x > p then (ls,x::rs) else (x::ls,rs);

qsort l = match l with | nil → nil
        | (x::xs) → let (ls,rs) = split (x,xs) in
                    append(qsort ls, x::(qsort rs));
```

With the evaluation-step metric we infer the following types.

$$\begin{array}{ll} \text{append:} & (L^{(8,0)}(int), L^{(0,0)}(int)) \xrightarrow{0,0} L^{(0,0)}(int) \\ \text{split:} & L^{(50,24)}(int) \xrightarrow{5,0} (L^{(34,24)}(int), L^{(26,24)}(int)) \\ \text{qsort:} & L^{(26,24)}(int) \xrightarrow{3,0} L^{(0,0)}(int) \end{array}$$

Thus *qsort* uses at most $3 + 14n + 12n^2$ evaluation steps. The function is typed resource-monomorphically in the recursive call *qsort rs* and resource-polymorphically in the recursive call *qsort ls*. The typing *qsort:* $L^{(34,24)}(int) \xrightarrow{3,0} L^{(8)}(int)$ is used there to cover the following costs of *append*.

[4] Note that these symbolic bounds are also part of the output of the analysis in our prototype implementation.

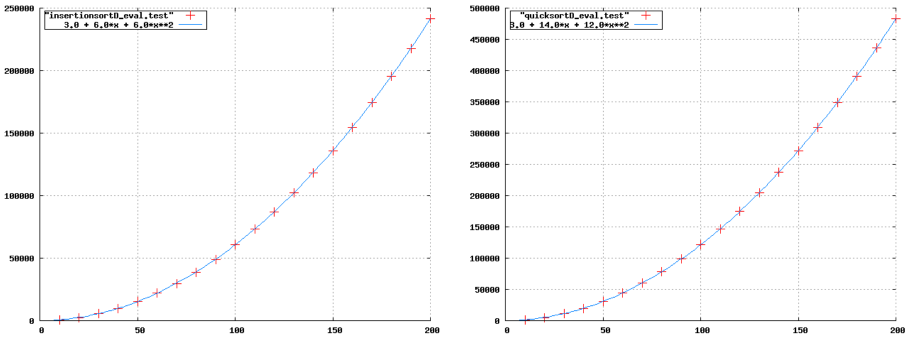

Fig. 4. The computed evaluation-step bound (blue line) compared to the actual worst-case number of evaluation-steps for sample inputs of various sizes (red crosses) used by *isort* (on the left) and *qsort* (on the right).

As the computed bounds indicate, insertion sort indeed admits a better worst-case behavior than quick sort. The reason is that there is an (expansive) call of *append* at each recursive call of *qsort*. Below is a tail-recursive version of quick sort that does not use *append*.

```
q_aux(l,acc) = match l with | nil → acc
                 | (x::xs) → let (ls,rs) = split (x,xs) in
                             let acc' = x::q_aux(rs,acc) in q_aux(ls,acc');
qsort2 l = q_aux(l,[]);
```

The prototype infers the following types.

$$\text{q\_aux:}\ \ (L^{(26,16)}(\mathit{int}), L^{(0,0)}(\mathit{int})) \xrightarrow{3.0} L^{(0,0)}(\mathit{int})$$
$$\text{qsort2:}\ L^{(26,16)}(\mathit{int}) \xrightarrow{7.0} L^{(0,0)}(\mathit{int})$$

The bound for *qsort2* is $7 + 18n + 8n^2$. It improves the bound of *qsort* in the quadratic part. The reduced potential in the second position of the type annotation of the argument corresponds directly to the costs for the calls of *append*. However, insertion sort has still a slightly better bound. Since *qsort2* is tail recursive, there is no need to use resource-polymorphic recursion in the type derivation.

Merge Sort The last function we implement is merge sort.

```
msplit l = match l with | nil → (nil,nil)
              | (x1::xs) → match xs with | nil → ([x1],nil)
                                | (x2::xs') → let (l1,l2) = msplit xs' in
                                                (x1::l1,x2::l2);
merge (l1,l2) = match l1 with | nil → l2
                  | (x::xs) → match l2 with | nil → (x::xs)
                                   | (y::ys) → if x<y then x::merge(xs,y::ys)
```

```
                                                    else y :: merge(x::xs,ys);
msort l = match l with | nil → nil
             | (x1 :: xs) → match xs with | nil → l
                               | (x2 :: xs') → let (l1,l2) = msplit l in
                                                 merge (msort l1, msort l2);
```

The following types are computed with the evaluation-step metric.

$$\begin{array}{l} \text{msplit: } L^{(46,0)}(\mathit{int}) \xrightarrow{25,0} (L^{(16,92)}(\mathit{int}), L^{(16,92)}(\mathit{int})) \\ \text{merge: } (L^{(16,0)}(\mathit{int}), L^{(16,0)}(\mathit{int})) \xrightarrow{3,0} L^{(0,0)}(\mathit{int}) \\ \text{msort: } L^{(0,92)}(\mathit{int}) \xrightarrow{5,0} L^{(0,0)}(\mathit{int}) \end{array}$$

The evaluation-step bound for *msort* is $5 - 46n + 46n^2$. In both recursive calls, the function is used resource-polymorphically with the alternate typing *msort:* $L^{(16,92)}(\mathit{int}) \xrightarrow{5,0} L^{(16,0)}(\mathit{int})$. Although our system cannot express an asymptotically tight $O(n \log n)$ bound for the function, it doubles the quadratic potential in the result of *msplit* and thus implicitly infers that *msplit* divides a list into two sublist of about equal length.

RAML vs. Knuth 1. Speed. Our prototype runs less then a second on a Macbook Pro with a 2.16 GHz Core 2 Duo to compute the bounds for all the above functions. Even though Knuth is known for his high productivity, such a performance seems to be hardly manually achievable. **2. Quality.** For sorting algorithms it is easy to identify inputs for which the worst-case run time behavior emerges. So we measured the actual worst-case behavior of the algorithms for several input sizes and compared it to the inferred bounds. Fig. 4 shows the results of these experiments for *qsort* and *isort*. The measured worst-case behavior of the functions matches exactly our computed bound. In fact, RAML computes tight evaluation-step bounds for *isort*, *qsort*, and *qsort2*. An automatic analysis can of course not always achieve the same accuracy as a careful manual analysis. In the case of RAML, we can only deal with polynomial bounds. Since the actual worst-case behavior of merge sort is $O(n \log n)$, the inferred quadratic bound is loose. **3. Scalability.** There is still a large number of polynomially bounded programs that cannot be analyzed in our system. On the other hand, a manual analysis is nearly impossible for large programs that are written in a high-level language. **4. Practicability.** First, a manual analysis of assembly code is tedious and error-prone. Secondly, it is time intensive and expensive. Thirdly, it has to be repeated after every change in the program which can lead to subtle errors due to false assumptions about the nature of the change. In contrast, the RAML analysis is proved to be sound and available at the touch of a button every time the program changed.

8 Related Work

Most closely related is the previous work on automatic amortized analysis [9, 1, 3, 5, 7] (see §1). This paper focuses on polymorphic recursion and is the first that investigates relations of the inferred bounds to non-terminating computations. A major conceptual innovation is the extension of the amortized method to polynomial bounds that we have introduced in a companion paper [9].

Other resource analyses that can in principle obtain polynomial bounds are approaches based on recurrences pioneered by Grobauer [13] and Flajolet [14]. In those systems, an a priori unknown resource bounding function is introduced for each function in the code; by a straightforward intraprocedural analysis a set of recurrence equations or inequations for these functions is then derived. A type-based extraction of such recurrences has been given in [15]. Even for relatively simple programs the resulting recurrences are quite complicated and difficult to solve with standard methods. In the COSTA project [16, 17] progress has been made with the solution of those recurrences. In an automatic complexity analysis for higher-order Nuprl terms Benzinger uses Mathematica to solve the generated recurrence equations [18]. Still, we find that amortization yields better results in cases where resource usage of intermediate functions depends on factors other than input size, e.g., sizes of partitions in quick sort. Also compositions of functions seem to be better dealt with by amortization.

A successful method to estimate time bounds for C++ procedures with loops and recursion was recently developed by Gulwani et al. [19–21] in the SPEED project. They annotate programs with counters and use automatic invariant discovery between their values using off-the-shelf program analysis tools which are based on abstract interpretation. A recent innovation for non-recursive programs is the combination of disjunctive invariant generation via abstract interpretation with proof rules that employ SMT-solvers [22]. In contrast to our method, these techniques can not fully automatically analyze iterations over data structures. Instead, the user needs to define numerical "quantitative functions". A methodological difference is that we infer (using linear programming) an abstract potential function which indirectly yields a resource-bounding function. The potential-based approach may be favorable in the presence of compositions and data scattered over different locations (partitions in quick sort). Moreover, our method infers tight bounds for functions like insertion sort that admit a worst-case time usage of the form $\sum_{1\leq i \leq n} i$. In contrast, [19] indicates that a nested loop on $1 \leq i \leq n$ and $1 \leq j \leq i$ is over-approximated with the bound n^2.

The examples from loc. cit. suggest that the two approaches are complementary in the sense that the method of Gulwani et al. works well for programs with little or no recursion but integrate interaction of linear arithmetic with loops. Our method, on the other hand, does not model the interaction of integer arithmetic with resource usage, but is particularly good for analyzing recursive programs involving inductive data types. As any type system, our approach is naturally compositional and lends itself to the smooth integration of components whose implementation is not available. Moreover, type derivations can be seen as certificates and can be automatically translated into formalized proofs in program logic [23].

Another related approach is the use of sized types [24–27] which provide a general framework to represent the size of the data in its type. Sized types are a very important concept and we also employ them indirectly. Our method adds a certain amount of data dependency and dispenses with the explicit manipulation of symbolic expressions in favour of numerical potential annotations.

Polynomial resource bounds have also been studied in [28] that addresses the derivation of polynomial size bounds for functions whose exact growth rate is polynomial. Besides this strong restriction, the efficiency of inference remains unclear.

9 Conclusion and Future Research

We have continued our work on automatic type-base amortized analysis for polynomial resource bounds.

To deal with the challenge of resource-polymorphic recursion we have introduced a new inference algorithm. It uses a special cost-free resource metric to compute alternate function types for recursive calls. The algorithm has been implemented and it has been shown by experiments that it efficiently computes types for interesting examples such as sorting algorithms.

To prove the non-trivial soundness of the algorithm for terminating and non-terminating evaluations we introduced a novel partial big-step operational semantics. It models non-termination with inductive rules.

Even though there are examples that the inference algorithm cannot handle we find it to be a good compromise between efficiency and performance. Therefore, our future research will focus mainly on conceptual extensions of the type system that will employ the same inference method. Most notably we plan an extension to mixed potential capable of inferring bounds like $n \cdot m$, an extension to recursion on non-inductive data like integers, and the integration of higher-order and polymorphism. We also investigate methods to derive non-polynomial bounds like 2^n and $n \log n$.

References

1. Hofmann, M., Jost, S.: Static Prediction of Heap Space Usage for First-Order Functional Programs. In: 30th ACM Symp. on Principles of Prog. Langs. (POPL'03). (2003) 185–197
2. Tarjan, R.E.: Amortized Computational Complexity. SIAM J. Algebraic Discrete Methods **6**(2) (1985) 306–318
3. Hofmann, M., Jost, S.: Type-Based Amortised Heap-Space Analysis. In: Prog. Langs. and Systems, 15th European Symp. on Prog. (ESOP'06). (2006) 22–37
4. Hofmann, M., Rodriguez, D.: Efficient Type-Checking for Amortised Heap-Space Analysis. In: 18th Conf. on Comp. Science Logic (CSL'09), LNCS (2009)
5. Jost, S., Loidl, H.W., Hammond, K., Scaife, N., Hofmann, M.: Carbon Credits for Resource-Bounded Computations using Amortised Analysis. In: 16th Intl. Symp. on Form. Meth. (FM'09). (2009) 354–369
6. Campbell, B.: Amortised Memory Analysis using the Depth of Data Structures. In: 18th Euro. Symp. on Prog. (ESOP'09). (2009) 190–204

7. Jost, S., Hammond, K., Loidl, H.W., Hofmann, M.: Static Determination of Quantitative Resource Usage for Higher-Order Programs. In: 37th ACM Symp. on Principles of Prog. Langs. (POPL'10). (2010) 223–236
8. Atkey, R.: Amortised Resource Analysis with Separation Logic. In: 19th Euro. Symp. on Prog. (ESOP'10). (2010) 85–103
9. Hoffmann, J., Hofmann, M.: Amortized Resource Analysis with Polynomial Potential. In: 19th Euro. Symp. on Prog. (ESOP'10). (2010) 287–306
10. Cousot, P., Cousot, R.: Inductive Definitions, Semantics and Abstract Interpretations. In: 19th ACM Symp. on Principles of Prog. Langs. (POPL '92). (1992) 83–94
11. Leroy, X.: Coinductive Big-Step Operational Semantics. In: 15th Euro. Symp. on Prog. (ESOP'06). (2006) 54–68
12. Knuth, D.E.: The Art of Computer Programming, Volume 1 (3rd ed.): Fundamental Algorithms. Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA (1997)
13. Grobauer, B.: Cost Recurrences for DML Programs. In: 6th Intl. Conf. on Funct. Prog. (ICFP'01). (2001) 253–264
14. Flajolet, P., Salvy, B., Zimmermann, P.: Automatic Average-case Analysis of Algorithms. Theoret. Comput. Sci. **79**(1) (1991) 37–109
15. Crary, K., Weirich, S.: Resource Bound Certification. In: 27th ACM Symp. on Principles of Prog. Langs. (POPL'00). (2000) 184–198
16. Albert, E., Arenas, P., Genaim, S., Puebla, G., Zanardini, D.: Cost Analysis of Java Bytecode. In: 16th Euro. Symp. on Prog. (ESOP'07). (2007) 157–172
17. Albert, E., Arenas, P., Genaim, S., Puebla, G.: Automatic Inference of Upper Bounds for Recurrence Relations in Cost Analysis. In: Static Analysis, 15th Intl. Symp. (SAS'08). (2008) 221–237
18. Benzinger, R.: Automated Higher-Order Complexity Analysis. Theor. Comput. Sci. **318**(1-2) (2004) 79–103
19. Gulwani, S., Mehra, K.K., Chilimbi, T.M.: SPEED: Precise and Efficient Static Estimation of Program Computational Complexity. In: 36th ACM Symp. on Principles of Prog. Langs. (POPL'09). (2009) 127–139
20. Gulavani, B.S., Gulwani, S.: A Numerical Abstract Domain Based on Expression Abstraction and Max Operator with Application in Timing Analysis. In: Comp. Aid. Verification, 20th Int. Conf. (CAV '08). (2008) 370–384
21. Gulwani, S., Jain, S., Koskinen, E.: Control-Flow Refinement and Progress Invariants for Bound Analysis. In: Conf. on Prog. Lang. Design and Impl. (PLDI'09). (2009) 375–385
22. Gulwani, S., Zuleger, F.: The Reachability-Bound Problem. In: Conf. on Prog. Lang. Design and Impl. (PLDI'10). (2010) 292–304
23. Beringer, L., Hofmann, M., Momigliano, A., Shkaravska, O.: Automatic Certification of Heap Consumption. In: Logic for Prog., AI, and Reasoning, 11th Int. Conf. (LPAR'04). (2004) 347–362
24. Hughes, J., Pareto, L., Sabry, A.: Proving the Correctness of Reactive Systems Using Sized Types. In: Symp. Princ. of Prog. Langs. (POPL'96). (1996) 410–423
25. Hughes, J., Pareto, L.: Recursion and Dynamic Data-structures in Bounded Space: Towards Embedded ML Programming. In: 4th Intl. Conf. on Funct. Prog. (ICFP'99). (1999) 70–81
26. Chin, W.N., Khoo, S.C.: Calculating Sized Types. High.-Ord. and Symb. Comp. **14**(2-3) (2001) 261–300

27. Chin, W.N., Khoo, S.C., Qin, S., Popeea, C., Nguyen, H.H.: Verifying Safety Policies with Size Properties and Alias Controls. In: Intl. Conf. on Software Eng. (ICSE'05). (2005) 186–195
28. Shkaravska, O., van Kesteren, R., van Eekelen, M.C.: Polynomial Size Analysis of First-Order Functions. In: Typed Lambda Calc. Apps. (TLCA'07). (2007) 351–365

A Soundness Proof

Theorem (Soundness) *Let $\mathcal{H} \vDash \mathcal{V}{:}\Gamma$ and let $\Gamma \; {}^{k}{\vdash}^{q}_{q'} \; e{:}A$.*

1. *If $\mathcal{V}, \mathcal{H} \vdash e \rightsquigarrow v, \mathcal{H}' \mid (p, p')$ then $p \leq \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + q$ and $p - p' \leq \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + q - (\Phi_{\mathcal{H}'}(v{:}A) + q')$.*
2. *If $\mathcal{V}, \mathcal{H} \vdash e \rightsquigarrow \mid p$ then $p \leq \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + q$.*

Proof of Part 1. We prove the statement simultaneously for all instantiations of the resource constants. It proceeds by induction on the derivation of the evaluation judgment $\mathcal{V}, \mathcal{H} \vdash e \rightsquigarrow v, \mathcal{H}' \mid (p, p')$.

Base Case In the base case the evaluation consists of a single step and e is a constant c, a variable x, a binary operation $x_1 \; \mathit{op} \; x_2$, a pair (x_1, x_2), the constant *nil*, or $\mathit{cons}(x_1, x_2)$. We show the argument for variables and concatenations only. The other cases are similar.

(E:Var) Assume that e is a variable x that has been evaluated with the rule E:Var. Assume first that $K^{\text{var}} \geq 0$. Then it follows by definition that $p = K^{\text{var}}$, $p' = 0$ and the type judgment $\Gamma \; {}^{k}{\vdash}^{q}_{q'} \; x{:}A$ has been derived by a single application of the rule T:Var. Thus we have $q \geq q' + K^{\text{var}}$ and therefore $p = K^{\text{var}} \leq q \leq \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + q$ and $p - p' = K^{\text{var}} \leq q - q' \leq \Phi_{\mathcal{V},\mathcal{H}}(\Gamma') + q - q' = \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + q - (\Phi_{\mathcal{H}'}(v{:}A) + q')$ (where $\Gamma = \Gamma', x{:}A$ and $\mathcal{H}' = \mathcal{H}$).

Assume now that $K^{\text{var}} < 0$. Then it follows by definition that $p = 0$, $p' = -K^{\text{var}}$. Thus $p = 0 \leq \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + q$. We have again that $q \geq q' + K^{\text{var}}$ and $p - p' = K^{\text{var}}$. Therefore the second part of the statement follows like in the case $K^{\text{var}} \geq 0$.

(E:Cons) If e has the form $\mathit{cons}(x_h, x_t)$ then it has been evaulated with the rule E:Cons. It follows by definition that $\mathcal{V}, \mathcal{H} \vdash \mathit{cons}(x_h, x_t) \rightsquigarrow l, \mathcal{H}[l \mapsto v'] \mid K^{\text{cons}}$, $x_h, x_t \in \mathrm{dom}(\mathcal{V})$, $v = (\mathcal{V}(x_h), \mathcal{V}(x_t))$, and $l \notin \mathrm{dom}(\mathcal{H})$. Thus

$$p = K^{\text{cons}} \text{ and } p' = 0 \tag{5}$$

or (if $K^{\text{cons}} < 0$)

$$p = 0 \text{ and } p' = -K^{\text{cons}} \tag{6}$$

The type judgment

$$\Gamma', x_h{:}A'_1, x_t{:}L^{\vec{r}}(A'_2) \; {}^{k}{\vdash}^{q}_{q'} \; \mathit{cons}(x_h, x_t){:}L^{\vec{s}}(A')$$

has been derived by a single application of the rule T:Cons and we have

$$q \geq q' + s_1 + K^{\text{cons}}, \vec{r} \geq \lhd(\vec{s}) \text{ and } A_i <: A \tag{7}$$

If $p = 0$ then $p \leq \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + q$ holds because of our implicit side condition $q \geq 0$. Otherwise we have $p = K^{\text{cons}} \leq q \leq \Phi_{\mathcal{V},\mathcal{H}}(\Gamma)c + q$.

From the definition of Φ and <: it follows that

$$s_1 + \Phi_{\mathcal{V},\mathcal{H}}(x_h{:}A'_1, x_t{:}L^{\vec{r}}(A'_2)) \geq \Phi_{\mathcal{H}[l\mapsto v']}(l : L^{\vec{s}}(A')) \tag{8}$$

Therefore

$$\begin{aligned}\Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + q &\geq \Phi_{\mathcal{V},\mathcal{H}}(x_h{:}A, x_t{:}L^{\vec{r}}(A')) + q \\ &\overset{(7)}{\geq} \Phi_{\mathcal{V},\mathcal{H}}(x_h{:}A, x_t{:}L^{\vec{r}}(A')) + q' + s_1 + K^{\text{cons}} \\ &\overset{(8)}{\geq} q' + K^{\text{cons}} + \Phi_{\mathcal{H}[l\mapsto v']}(l : L^{\vec{s}}(A'))\end{aligned}$$

and thus $\Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + q - (\Phi_{\mathcal{H}'}(v{:}A) + q') \geq K^{\text{cons}} = p - p'$.

Induction Case If the evaluation needs more then one step then e is a pattern match, a function application, a conditional, or a let expression.

(E:MatP) Assume that e is a pattern match for a pair that has the form *match* x *with* $(x_1, x_2) \to e'$. Then the rule E:MatP has been used at the rude of the derivation of the evaluation judgment. Therefore we have $\mathcal{V}(x) = (v_1, v_2)$ and $\mathcal{V}', \mathcal{H} \vdash e' \rightsquigarrow v, \mathcal{H}' \mid (r, r')$ for $\mathcal{V}[x_1 \mapsto v_1, x_2 \mapsto v_2]$ and some r, r' with

$$(p, p') = K_1^{\text{matP}} \cdot (r, r') \cdot K_2^{\text{matP}} \tag{9}$$

Similarly, the type judgment for e has been derived by an application of the rule T:MatP and thus $\Gamma = \Gamma', x{:}B,\quad B = (B_1, B_2),\quad \Gamma', x_1{:}B_1, x_2{:}B_2 \;{}^{k}\!\vdash^{s}_{s'}\; e : A$, and

$$q \geq s + K_1^{\text{matP}} \text{ and } s' - K_2^{\text{matP}} \geq q' \geq 0 \tag{10}$$

for some B_1, B_2, s, s'. Since $\mathcal{H} \vDash \mathcal{V}' : \Gamma', x_1{:}B_1, x_2{:}B_2$ we can apply the induction hypothesis and $\Phi_{\mathcal{V}',\mathcal{H}}(\Gamma', x_1{:}B_1, x_2{:}B_2) = \Phi_{\mathcal{V},\mathcal{H}}(\Gamma)$ to derive

$$r \leq \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + s \tag{11}$$

$$r - r' \leq \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + s - (\Phi_{\mathcal{H}'}(v{:}A) + s') \tag{12}$$

Let

$$(u, u') = K_1^{\text{matP}} \cdot (\Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + s, \Phi_{\mathcal{H}'}(v{:}A) + s') \cdot K_2^{\text{matP}} \tag{13}$$

Per definition and from (10) it follows that $u = \max(0, s + K_1^{\text{matP}} + \Phi_{\mathcal{V},\mathcal{H}}(\Gamma))$ (recall that K_1^{matP} might be negative). From Prop. 1 applied to (11), (13) and (9) we derive $u \geq p$. If $s + K_1^{\text{matP}} + \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) \leq 0$ then $u = p = 0$ and $q + \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) \geq p$ trivially holds. If $s + K_1^{\text{matP}} + \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) > 0$ then it follows from (10) that

$$q + \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) \geq s + K_1^{\text{matP}} + \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) = u \geq p$$

Similarly, we apply Prop. 1 to (9) and use (12) and (10) to see that

$$\begin{aligned}p - p' &= r - r' + K_1^{\text{matP}} + K_2^{\text{matP}} \\ &\leq \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + s - (\Phi_{\mathcal{H}'}(v{:}A) + s') + K_1^{\text{matP}} + K_2^{\text{matP}} \\ &\leq \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + (s + K_1^{\text{matP}}) - (\Phi_{\mathcal{H}'}(v{:}A) + (s' - K_2^{\text{matP}}) \\ &\leq \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + q - (\Phi_{\mathcal{H}'}(v{:}A) + q')\end{aligned}$$

The cases where the derivation of the evaluation judgment of e ends with an application of the rules E:CondT, E:CondT, or E:MatN are similar to the case E:MatP.

(E:MatC) Let e be a pattern match of the form

$$\mathit{match}\ x\ \mathit{with} \mid \mathit{nil} \to e_1 \mid \mathit{cons}(x_h, x_t) \to e_2$$

whose evaluation ends with an application of E:MatC. Then $\mathcal{V}(x) = l$, $\mathcal{H}(l) = (v_h, v_t)$, and $\mathcal{V}', \mathcal{H} \vdash e_2 \rightsquigarrow v, \mathcal{H}' \mid (r, r')$ for $\mathcal{V}' = \mathcal{V}[x_h \mapsto v_h, x_t \mapsto v_t]$ and some r, r' with

$$(p, p') = K_1^{\mathrm{matC}} \cdot (r, r') \cdot K_2^{\mathrm{matC}} \tag{14}$$

Furthermore, the derivation of $\varGamma \;{}^{k}{\vdash}^{q}_{q'}\; e{:}A$ ends with an application of T:MatL and hence we have $\varGamma = \varGamma', x{:}L^{\vec{t}}(B)$, $\;\varGamma', x_h{:}B, x_t{:}L^{\lhd(\vec{t})}(B) \;{}^{k}{\vdash}^{s}_{s'}\; e_2 : A_2$, $\;A_2 {<:} A$ and

$$q \geq s + K_1^{\mathrm{matC}} - t_1 \text{ and } s' - K_2^{\mathrm{matC}} \geq q' \geq 0 \tag{15}$$

It is true that $\varPhi_{\mathcal{H}}(v{:}L^{\vec{t}}(B)) = t_1 + \varPhi_{\mathcal{H}}(v_h{:}B) + \varPhi_{\mathcal{H}}(v_t{:}L^{\lhd(\vec{t})}(B))$ and therefore

$$\varPhi_{\mathcal{H},\mathcal{V}}(\varGamma) = t_1 + \varPhi_{\mathcal{H},\mathcal{V}'}(\varGamma', x_h{:}B, x_t{:}L^{\lhd(\vec{t})}(B)) \tag{16}$$

Since $\mathcal{H} \vDash \mathcal{V}' : \varGamma', x_h{:}B, x_t{:}L^{\lhd(\vec{t})}(B)$ we can apply the induction hypothisis to $\mathcal{V}', \mathcal{H} \vdash e_2 \rightsquigarrow v, \mathcal{H}' \mid (r, r')$ and obtain (with (16))

$$r \leq \varPhi_{\mathcal{V},\mathcal{H}}(\varGamma) - t_1 + s \tag{17}$$

$$r - r' \leq \varPhi_{\mathcal{V},\mathcal{H}}(\varGamma) - t_1 + s - (\varPhi_{\mathcal{H}'}(v{:}A_2) + s') \tag{18}$$

Note that $\varPhi_{\mathcal{V},\mathcal{H}}(\varGamma) - t_1 \geq 0$ and let

$$(u, u') = K_1^{\mathrm{matC}} \cdot (\varPhi_{\mathcal{V},\mathcal{H}}(\varGamma) - t_1 + s, \varPhi_{\mathcal{H}'}(v{:}A_2) + s') \cdot K_2^{\mathrm{matC}} \tag{19}$$

Per definition and from (15) it follows that $u = \max(0, \varPhi_{\mathcal{V},\mathcal{H}}(\varGamma) - t_1 + s + K_1^{\mathrm{matC}})$. From Prop. 1 applied to (17), (19) and (14) we derive $u \geq p$. If $\varPhi_{\mathcal{V},\mathcal{H}}(\varGamma) - t_1 + s + K_1^{\mathrm{matC}} \leq 0$ then $u = p = 0$ and $q + \varPhi_{\mathcal{V},\mathcal{H}}(\varGamma) \geq p$ trivially holds. If $\varPhi_{\mathcal{V},\mathcal{H}}(\varGamma) - t_1 + s + K_1^{\mathrm{matC}} > 0$ then it follows from (15) that

$$q + \varPhi_{\mathcal{V},\mathcal{H}}(\varGamma) \geq \varPhi_{\mathcal{V},\mathcal{H}}(\varGamma) - t_1 + s + K_1^{\mathrm{matC}} = u \geq p$$

Finally, we apply Prop. 1 to (14) to see that

$$\begin{aligned}
p - p' &= r - r' + K_1^{\mathrm{matC}} + K_2^{\mathrm{matC}} \\
&\leq \varPhi_{\mathcal{V},\mathcal{H}}(\varGamma) - t_1 + s - (\varPhi_{\mathcal{H}'}(v{:}A_2) + s') + K_1^{\mathrm{matC}} + K_2^{\mathrm{matC}} \\
&\overset{(18)}{\leq} \varPhi_{\mathcal{V},\mathcal{H}}(\varGamma) - t_1 + s - (\varPhi_{\mathcal{H}'}(v{:}A) + s') + K_1^{\mathrm{matC}} + K_2^{\mathrm{matC}} \\
&= \varPhi_{\mathcal{V},\mathcal{H}}(\varGamma) + (s + K_1^{\mathrm{matC}} - t_1) - (\varPhi_{\mathcal{H}'}(v{:}A) + (s' - K_2^{\mathrm{matC}})) \\
&\overset{(15)}{\leq} \varPhi_{\mathcal{V},\mathcal{H}}(\varGamma) + q - (\varPhi_{\mathcal{H}'}(v{:}A) + q')
\end{aligned}$$

(E:FunApp) Assume that e is a function application of the form $f(x)$. The evaluation of for e then ends with an application of the rule E:FunApp. Thus we have $\mathcal{V}(x) = v'$ and $[y_f \mapsto v'], \mathcal{H} \vdash e_f \leadsto v, \mathcal{H}' \mid (r, r')$ for some r, r' with

$$K_1^{\mathrm{app}} \cdot (r, r') \cdot K_2^{\mathrm{app}} \tag{20}$$

($k = 1$) If $k = 1$ then the derivation of the type judgment for e ends with an application of T:FunApp1. Therefore it is true that $\Gamma = \Gamma', x{:}B$, $\Sigma(f) = B \xrightarrow{s/s'} A$, and for some $c \in \mathbb{Q}^+$

$$q = s + c + K_1^{\mathrm{app}} \text{ and } q' = s' + c - K_2^{\mathrm{app}} \tag{21}$$

In order to apply the induction hypothesis to the evaluation of the function body e_f we recall from the definition of a well-formed program that $\Sigma(f) = B \xrightarrow{s/s'} A$ means that $y_f{:}B \overset{s}{\underset{s'}{\vdash}} e_f{:}A$. Since $\mathcal{H} \vDash \mathcal{V}{:}\Gamma', x{:}B$ and $\mathcal{V}(x) = v'$ it follows $\mathcal{H} \vDash [y_f \mapsto v'] : y_f{:}B$. We obtain by induction that

$$r \le \Phi_{[y_f \mapsto v'],\mathcal{H}}(y_f{:}B) + s \tag{22}$$

$$r - r' \le \Phi_{[y_f \mapsto v'],\mathcal{H}}(y_f{:}B) + s - (\Phi_{\mathcal{H}'}(v{:}A) + s') \tag{23}$$

Now everything is in place to proceed as in the case of E:MatP. Let

$$(u, u') = K_1^{\mathrm{app}} \cdot (\Phi_{[y_f \mapsto v'],\mathcal{H}}(y_f{:}B) + s, \Phi_{\mathcal{H}'}(v{:}A) + s') \cdot K_2^{\mathrm{app}} \tag{24}$$

to see that $p \le u = \max(0, K_1^{\mathrm{app}} + \Phi_{[y_f \mapsto v'],\mathcal{H}}(y_f{:}B) + s)$. Furthermore we have $\Phi_{[y_f \mapsto v'],\mathcal{H}}(y_f{:}B) = \Phi_{\mathcal{V},\mathcal{H}}(x{:}B) \le \Phi_{\mathcal{V},\mathcal{H}}(\Gamma)$ and with (21) it follows that $p \le \max(0, q - c + \Phi_{\mathcal{V},\mathcal{H}}(\Gamma)) \le q + \Phi_{\mathcal{V},\mathcal{H}}(\Gamma)$.

For the second part for the statement observe that

$$\begin{aligned}
p - p' &= r - r' + K_1^{\mathrm{app}} + K_2^{\mathrm{app}} \\
&\overset{(23)}{\le} \Phi_{[y_f \mapsto v'],\mathcal{H}}(y_f{:}B) + s - (\Phi_{\mathcal{H}'}(v{:}A) + s') + K_1^{\mathrm{app}} + K_2^{\mathrm{app}} \\
&\le \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + s - (\Phi_{\mathcal{H}'}(v{:}A) + s') + K_1^{\mathrm{app}} + K_2^{\mathrm{app}} \\
&= \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + s + K_1^{\mathrm{app}} - (\Phi_{\mathcal{H}'}(v{:}A) + s' - K_2^{\mathrm{app}}) \\
&\overset{(21)}{=} \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + q - (\Phi_{\mathcal{H}'}(v{:}A) + q')
\end{aligned}$$

($k > 1$) If $k > 1$ then the derivation of the type judgment for e ends with an application of T:FunApp. Therefore $\Sigma_{cf}; y_f{:}B_{cf} \;{}^{cf(k-1)}\!\overset{s_{cf}}{\underset{s_{cf}}{\vdash}} e_f{:}A_{cf}$, $\Gamma = \Gamma', x{:}B$, $\Sigma(f) = B' \xrightarrow{s/s'} A'$, and for some $c \in \mathbb{Q}^+$

$$q = s + s_{cf} + c + K_1^{\mathrm{app}} \text{ and } q' = s' + s'_{cf} + c - K_2^{\mathrm{app}} \tag{25}$$

$$\curlyvee(A \mid A', A_{cf}) \text{ and } \curlyvee(B \mid B', B_{cf}) \tag{26}$$

In order to apply the induction hypothesis to the evaluation of the function body e_f we recall from the definition of a well-formed program that $\Sigma(f) = B' \xrightarrow{s/s'} A'$

means that $y_f{:}B' \vdash^{s}_{s'} e_f{:}A'$. Since $\mathcal{H} \vDash \mathcal{V} : \Gamma', x{:}B'$ and $\mathcal{V}(x) = v'$ it follows $\mathcal{H} \vDash [y_f \mapsto v'] : y_f{:}B'$. We conclude by induction that

$$r \leq \Phi_{[y_f \mapsto v'],\mathcal{H}}(y_f{:}B') + s \tag{27}$$

$$r - r' \leq \Phi_{[y_f \mapsto v'],\mathcal{H}}(y_f{:}B') + s - (\Phi_{\mathcal{H}'}(v{:}A') + s') \tag{28}$$

Similarly, we use the induction hypothesis a second time for the *cost-free* instantiation of the resource constants. It is the case that $\Sigma_{cf}; x{:}B_{cf}\ {}^{cf\,(k\,-\,1)}\!\vdash^{s_{cf}}_{s_{cf}} e_f{:}A_{cf}$ and $[y_f \mapsto v'], \mathcal{H} \vdash e_f \rightsquigarrow v, \mathcal{H}' \mid (0,0)$ since all resource constants are zero. Thus

$$0 \leq \Phi_{[y_f \mapsto v'],\mathcal{H}}(y_f{:}B_{cf}) + s_{cf} - (\Phi_{\mathcal{H}'}(v{:}A_{cf}) + s'_{cf}) \tag{29}$$

Now consider

$$(u, u') {=} K_1^{\mathrm{app}} \cdot (\Phi_{[y_f \mapsto v'],\mathcal{H}}(y_f{:}B) + s, \Phi_{\mathcal{H}'}(v{:}A) + s') \cdot K_2^{\mathrm{app}} \tag{30}$$

and verify that $p \leq u = \max(0, K_1^{\mathrm{app}} + \Phi_{[y_f \mapsto v'],\mathcal{H}}(y_f{:}B') + s)$. It follows from (26) that $\Phi_{[y_f \mapsto v'],\mathcal{H}}(y_f{:}B') \leq \Phi_{\mathcal{V},\mathcal{H}}(x{:}B) \leq \Phi_{\mathcal{V},\mathcal{H}}(\Gamma)$ and thus with (25) $p \leq \max(0, q - c + \Phi_{\mathcal{V},\mathcal{H}}(\Gamma)) \leq q + \Phi_{\mathcal{V},\mathcal{H}}(\Gamma)$.

The second part is proved as follows.

$$\begin{aligned}
p - p' \ &= \ r - r' + K_1^{\mathrm{app}} + K_2^{\mathrm{app}} \\
&\overset{(28)}{\leq} \ \Phi_{[y_f \mapsto v'],\mathcal{H}}(y_f{:}B') + s - (\Phi_{\mathcal{H}'}(v{:}A') + s') + K_1^{\mathrm{app}} + K_2^{\mathrm{app}} \\
&\overset{(29)}{\leq} \ \Phi_{[y_f \mapsto v'],\mathcal{H}}(y_f{:}B') + s - (\Phi_{\mathcal{H}'}(v{:}A') + s') + K_1^{\mathrm{app}} + K_2^{\mathrm{app}} \\
&\qquad + \Phi_{[y_f \mapsto v'],\mathcal{H}}(y_f{:}B_{cf}) + s_{cf} - (\Phi_{\mathcal{H}'}(v{:}A_{cf}) + s'_{cf}) \\
&\overset{(26)}{=} \ \Phi_{[y_f \mapsto v'],\mathcal{H}}(y_f{:}B) + s + s_{cf} - (\Phi_{\mathcal{H}'}(v{:}A) + s' + s'_{cf}) + K_1^{\mathrm{app}} + K_2^{\mathrm{app}} \\
&\leq \ \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + s + s_{cf} + K_1^{\mathrm{app}} - (\Phi_{\mathcal{H}'}(v{:}A) + s' + s'_{cf} - K_2^{\mathrm{app}}) \\
&\overset{(25)}{=} \ \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + q - (\Phi_{\mathcal{H}'}(v{:}A) + q')
\end{aligned}$$

(E:Let) Let e be a let expression of the from *let* $x = e_1$ *in* e_2 that has eventually been evaluated with the rule E:Let. Then it follows that $\mathcal{V}, \mathcal{H} \vdash e_1 \rightsquigarrow v_1, \mathcal{H}_1 \mid (r, r')$ and $\mathcal{V}', \mathcal{H}_1 \vdash e_2 \rightsquigarrow v_2, \mathcal{H}_2 \mid (t, t')$ for $\mathcal{V}' = \mathcal{V}[x \mapsto v_1]$ and r, r', t, t' with

$$(p, p') = K_1^{\mathrm{let}} \cdot (r, r') \cdot K_2^{\mathrm{let}} \cdot (t, t') \cdot K_3^{\mathrm{let}} \tag{31}$$

The derivation of the type judgment for e ends with an application of T:Let. Hence $\Gamma = \Gamma_1, \Gamma_2, \Delta, \quad \Gamma_1, \Delta_1\ {}^{k}\!\vdash^{s_1}_{s'_1} e_1 : B,\ \Gamma_2, \Delta_2, x{:}B\ {}^{k}\!\vdash^{s_2}_{s'_2} e_2 : A$ and

$$\curlyvee(\Delta \mid \Delta_1, \Delta_2) \text{ and } \mathrm{Var}(\Gamma_1) \cap \mathrm{Var}(\Gamma_2) = \emptyset \tag{32}$$

$$q \geq s_1 + K_1^{\mathrm{let}} \tag{33}$$

$$s'_1 - K_2^{\mathrm{let}} \geq s_2 \tag{34}$$

$$s'_2 - K_3^{\mathrm{let}} \geq q' \tag{35}$$

It follows from (32) that

$$\Phi_{\mathcal{V},\mathcal{H}}(\Gamma) = \Phi_{\mathcal{V},\mathcal{H}}(\Gamma_1, \Delta_1) + \Phi_{\mathcal{V},\mathcal{H}}(\Gamma_2, \Delta_2) \tag{36}$$

Since $\mathcal{H} \vDash \mathcal{V} : \Gamma$ we have also $\mathcal{H} \vDash \mathcal{V} : \Gamma_1, \Delta_1$ and can thus apply the induction hypothesis for the evaluation judgment of e_1 to derive

$$r \leq \Phi_{\mathcal{V},\mathcal{H}}(\Gamma_1, \Delta_1) + s_1 \tag{37}$$
$$r - r' \leq \Phi_{\mathcal{V},\mathcal{H}}(\Gamma_1, \Delta_1) + s_1 - (\Phi_{\mathcal{H}_1}(v_1{:}B) + s'_1) \tag{38}$$

Form Thm. 2 it follows that $\mathcal{H}_2 \vDash \mathcal{V}' : \Gamma_2, \Delta_2, x{:}B$ and thus again by induction

$$t \leq \Phi_{\mathcal{V}',\mathcal{H}_1}(\Gamma_2, \Delta_2, x{:}B) + s_2 \tag{39}$$
$$t - t' \leq \Phi_{\mathcal{V}',\mathcal{H}_1}(\Gamma_2, \Delta_2, x{:}B) + s_2 - (\Phi_{\mathcal{H}_2}(v_2{:}A) + s'_2) \tag{40}$$

Now let

$$\begin{aligned}(u, u') = K_1^{\text{let}} \cdot (\Phi_{\mathcal{V},\mathcal{H}}(\Gamma_1, \Delta_1) + s_1, \Phi_{\mathcal{H}_1}(v_1{:}B) + s'_1) \cdot K_2^{\text{let}} \cdot \\ (\Phi_{\mathcal{V}',\mathcal{H}_1}(\Gamma_2, \Delta_2, x{:}B) + s_2, \Phi_{\mathcal{H}_2}(v_2{:}A) + s'_2) \cdot K_3^{\text{let}}\end{aligned}$$

Then it follows that

$$\begin{aligned}(u, u') &\overset{(34,35)}{=} K_1^{\text{let}} \cdot (\Phi_{\mathcal{V},\mathcal{H}}(\Gamma_1, \Delta_1) + s_1, \Phi_{\mathcal{H}_1}(v_1{:}B) + s'_1 - K_2^{\text{let}}) \cdot \\ &\qquad (\Phi_{\mathcal{V}',\mathcal{H}_1}(\Gamma_2, \Delta_2, x{:}B) + s_2, \Phi_{\mathcal{H}_2}(v_2{:}A) + s'_2 - K_3^{\text{let}}) \\ &= K_1^{\text{let}} \cdot (v + \Phi_{\mathcal{V},\mathcal{H}}(\Gamma_1, \Delta_1) + s_1, v')\end{aligned}$$

for some $v, v' \in \mathbb{Q}^+$ with

$$\begin{aligned}v &\leq \Phi_{\mathcal{V}',\mathcal{H}_1}(\Gamma_2, \Delta_2, x{:}B) + s_2 - (\Phi_{\mathcal{H}_1}(v_1{:}B) + s'_1 - K_2^{\text{let}}) \\ &\overset{(34)}{\leq} \Phi_{\mathcal{V}',\mathcal{H}_1}(\Gamma_2, \Delta_2, x{:}B)\end{aligned}$$

and thus

$$\begin{aligned}u &\leq \max(0, \Phi_{\mathcal{V},\mathcal{H}}(\Gamma_1, \Delta_1) + \Phi_{\mathcal{V}',\mathcal{H}_1}(\Gamma_2, \Delta_2) + s_1 + K_1^{\text{let}}) \\ &\overset{(\mathit{Lem.}\ 4)}{\leq} \max(0, \Phi_{\mathcal{V},\mathcal{H}}(\Gamma_1, \Delta_1) + \Phi_{\mathcal{V},\mathcal{H}}(\Gamma_2, \Delta_2) + s_1 + K_1^{\text{let}}) \\ &\overset{(32,33)}{\leq} \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + q\end{aligned}$$

Finally, it follows with Prop. 1 applied to (37), (39), and (31) that $u \geq p$.

For the second part of the statement we apply Prop. 1 to (31) to derive

$$\begin{array}{rcl}
p - p' & = & r - r' + t - t' + K^{\text{let}}_1 + K^{\text{let}}_2 + K^{\text{let}}_3 \\
& \overset{(40,38)}{\leq} & \Phi_{\mathcal{V},\mathcal{H}}(\Gamma_1, \Delta_1) + s_1 - (\Phi_{\mathcal{H}_1}(v_1{:}B) + s'_1) + \ \Phi_{\mathcal{V}',\mathcal{H}_1}(\Gamma_2, \Delta_2, x{:}B) \\
& & + \, s_2 - (\Phi_{\mathcal{H}_2}(v_2{:}A) + s'_2) + \ K^{\text{let}}_1 + K^{\text{let}}_2 + K^{\text{let}}_3 \\
& = & (\Phi_{\mathcal{V},\mathcal{H}}(\Gamma_1, \Delta_1) + \Phi_{\mathcal{V}',\mathcal{H}_1}(\Gamma_2, \Delta_2) + s_1) \\
& & + \, (s_2 + K^{\text{let}}_2 - s'_1) - (\Phi_{\mathcal{H}_2}(v_2{:}A) + s'_2) + K^{\text{let}}_1 + K^{\text{let}}_3 \\
& \overset{(34)}{\leq} & \Phi_{\mathcal{V},\mathcal{H}}(\Gamma_1, \Delta_1) + \Phi_{\mathcal{V}',\mathcal{H}_1}(\Gamma_2, \Delta_2) + s_1 - (\Phi_{\mathcal{H}_2}(v_2{:}A){+}s'_2){+}K^{\text{let}}_1{+}K^{\text{let}}_3 \\
& \overset{(L.\ 4)}{\leq} & \Phi_{\mathcal{V},\mathcal{H}}(\Gamma_1, \Delta_1) + \Phi_{\mathcal{V},\mathcal{H}}(\Gamma_2, \Delta_2) + s_1 - (\Phi_{\mathcal{H}_2}(v_2{:}A) + s'_2){+}K^{\text{let}}_1{+}K^{\text{let}}_3 \\
& \overset{(32)}{=} & \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + s_1 + K^{\text{let}}_1 - (\Phi_{\mathcal{H}_2}(v_2{:}A) + s'_2 - K^{\text{let}}_3) \\
& \overset{(33,35)}{\leq} & \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + q - (\Phi_{\mathcal{H}_2}(v_2{:}A) + q')
\end{array}$$

Proof of Part 2. The proof of part 2 is similar but simpler than the proof of part 1. However, it uses part 1 in the case of the rule P:Let2. As in part 1 we prove the statement by induction on the derivation of the partial evaluation judgment $\mathcal{V}, \mathcal{H} \vdash e \rightsquigarrow \mid p$.

We only show some cases to convince the reader that everything is analogue to the above induction.

(P:Var) Assume that e is a variable x that has been evaulated with a single application of the rule P:Var. Then it follows by definition that $p = \max(K^{\text{var}}, 0)$ and the type judgment $\Gamma \ {}^{k}{\vdash}^{q}_{q'} \ x{:}A$ has been derived by a single application of the rule T:Var. Thus we have $q \geq q' + K^{\text{var}}$ and therefore $p \leq q \leq \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + q$.

(P:Cons) If e has the form $\mathit{cons}(x_h, x_t)$ then it has been evaulated with the rule P:Cons. It follows by definition that $p = \max(K^{\text{cons}}, 0)$. Furthermore, the type judgment $\Gamma \ {}^{k}{\vdash}^{q}_{q'} \ e{:}A$ has been derived by a single application of the rule T:Cons. Thus $q \geq s_1 + K^{\text{cons}}$ and hence $\Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + q \geq q \geq p$.

(P:MatC) Let e be a pattern match of the form

$$\mathit{match}\ x\ \mathit{with} \mid \mathit{nil} \to e_1 \mid \mathit{cons}(x_h, x_t) \to e_2$$

whose evaluation ends with an application of P:MatC. Then $\mathcal{V}(x) = l$, $\mathcal{H}(l) = (v_h, v_t)$, and $\mathcal{V}', \mathcal{H} \vdash e_2 \rightsquigarrow \mid r$ for $\mathcal{V}' = \mathcal{V}[x_h \mapsto v_h, x_t \mapsto v_t]$ and some r with

$$p = \max(K^{\text{matC}}_1 + r, 0)$$

Furthermore, the derivation of $\Gamma \ {}^{k}{\vdash}^{q}_{q'} \ e{:}A$ ends with an application of T:MatL and hence we have $\Gamma = \Gamma', x{:}L^{\vec{t}}(B)$, $\ \Gamma', x_h{:}B, x_t{:}L^{\lhd(\vec{t})}(B) \ {}^{k}{\vdash}^{s}_{s'} \ e_2 : A_2$, and

$$q \geq s + K^{\text{matC}}_1 - t_1$$

It is true that $\Phi_{\mathcal{H}}(v{:}L^{\vec{t}}(B)) = t_1 + \Phi_{\mathcal{H}}(v_h{:}B) + \Phi_{\mathcal{H}}(v_t{:}L^{\lhd(\vec{t})}(B))$ and therefore

$$\Phi_{\mathcal{H},\mathcal{V}}(\Gamma) = t_1 + \Phi_{\mathcal{H},\mathcal{V}'}(\Gamma', x_h{:}B, x_t{:}L^{\lhd(\vec{t})}(B)) \tag{41}$$

Since $\mathcal{H} \vDash \mathcal{V}' : \Gamma', x_h{:}B, x_t{:}L^{\lhd(\vec{t})}(B)$ we can apply the induction hypothisis to $\mathcal{V}', \mathcal{H} \vdash e_2 \rightsquigarrow \mid r$ and obtain (with (41))

$$r \leq \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) - t_1 + s$$

If $p \geq 0$ we have thus $p = K_1^{\text{matC}} + r \leq K_1^{\text{matC}} + \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) - t_1 + s \leq \Phi_{\mathcal{V},\mathcal{H}}(\Gamma) + q$. If $p = 0$ then the statement follows since $q \geq 0$.

(P:Let2) Let e be a let expression of the from *let* $x = e_1$ *in* e_2 that has finally been evaluated with the rule P:Let2. Then it follows that $\mathcal{V}, \mathcal{H} \vdash e_1 \rightsquigarrow v_1, \mathcal{H}_1 \mid (r, r')$ and $\mathcal{V}', \mathcal{H}_1 \vdash e_2 \rightsquigarrow \mid t$ for $\mathcal{V}' = \mathcal{V}[x \mapsto v_1]$ and r, r', t with

$$(p, p') = K_1^{\text{let}} \cdot (r, r') \cdot K_2^{\text{let}} \cdot (t, 0) \tag{42}$$

The derivation of the type judgment for e ends with an application of T:Let. Hence $\Gamma = \Gamma_1, \Gamma_2, \Delta$, $\Gamma_1, \Delta_1 \ {}^{k}\!\!\vdash^{s_1}_{s'_1} e_1 : B$, $\ \Gamma_2, \Delta_2, x{:}B \ {}^{k}\!\!\vdash^{s_2}_{s'_2} e_2 : A$ and

$$\curlyvee(\Delta \mid \Delta_1, \Delta_2) \text{ and } \mathrm{Var}(\Gamma_1) \cap \mathrm{Var}(\Gamma_2) = \emptyset \tag{43}$$

$$q \geq s_1 + K_1^{\text{let}} \tag{44}$$

$$s'_1 - K_2^{\text{let}} \geq s_2 \tag{45}$$

It follows from (43) that

$$\Phi_{\mathcal{V},\mathcal{H}}(\Gamma) = \Phi_{\mathcal{V},\mathcal{H}}(\Gamma_1, \Delta_1) + \Phi_{\mathcal{V},\mathcal{H}}(\Gamma_2, \Delta_2) \tag{46}$$

Since $\mathcal{H} \vDash \mathcal{V} : \Gamma$ we have also $\mathcal{H} \vDash \mathcal{V} : \Gamma_1, \Delta_1$ and can thus apply part 1 of Thm. 5 to the evaluation judgment of e_1 to derive

$$r \leq \Phi_{\mathcal{V},\mathcal{H}}(\Gamma_1, \Delta_1) + s_1 \tag{47}$$

$$r - r' \leq \Phi_{\mathcal{V},\mathcal{H}}(\Gamma_1, \Delta_1) + s_1 - (\Phi_{\mathcal{H}_1}(v_1{:}B) + s'_1) \tag{48}$$

Form Thm. 2 it follows that $\mathcal{H}_2 \vDash \mathcal{V}' : \Gamma_2, \Delta_2, x{:}B$ and we can thus apply the induction hypothesis to the partial evaluation of e_2. We obtain

$$t \leq \Phi_{\mathcal{V}',\mathcal{H}_1}(\Gamma_2, \Delta_2, x{:}B) + s_2 \tag{49}$$

Now let

$$\begin{aligned}(u, u') = K_1^{\text{let}} \cdot (\Phi_{\mathcal{V},\mathcal{H}}(\Gamma_1, \Delta_1) + s_1, \Phi_{\mathcal{H}_1}(v_1{:}B) + s'_1)\\ \cdot K_2^{\text{let}} \cdot (\Phi_{\mathcal{V}',\mathcal{H}_1}(\Gamma_2, \Delta_2, x{:}B) + s_2, 0)\end{aligned}$$

Then it follows that

$$\begin{aligned}(u, u') &= K_1^{\text{let}} \cdot (\Phi_{\mathcal{V},\mathcal{H}}(\Gamma_1, \Delta_1) + s_1, \Phi_{\mathcal{H}_1}(v_1{:}B) + s'_1) \cdot K_2^{\text{let}} \cdot \\ &\quad (\Phi_{\mathcal{V}',\mathcal{H}_1}(\Gamma_2, \Delta_2, x{:}B) + s_2, 0) \\ &\overset{(45)}{=} K_1^{\text{let}} \cdot (\Phi_{\mathcal{V},\mathcal{H}}(\Gamma_1, \Delta_1) + s_1, \Phi_{\mathcal{H}_1}(v_1{:}B) + s'_1 - K_2^{\text{let}}) \cdot \\ &\quad (\Phi_{\mathcal{V}',\mathcal{H}_1}(\Gamma_2, \Delta_2, x{:}B) + s_2, 0) \\ &= K_1^{\text{let}} \cdot (v + \Phi_{\mathcal{V},\mathcal{H}}(\Gamma_1, \Delta_1) + s_1, v')\end{aligned}$$

for some $v, v' \in \mathbb{Q}^+$ with

$$\begin{array}{rl} v & \leq \mathit{\Phi}_{\mathcal{V}',\mathcal{H}_1}(\mathit{\Gamma}_2, \mathit{\Delta}_2, x{:}B) + s_2 - (\mathit{\Phi}_{\mathcal{H}_1}(v_1{:}B) + s'_1 - K^{\text{let}}_2) \\ & \overset{(45)}{\leq} \mathit{\Phi}_{\mathcal{V}',\mathcal{H}_1}(\mathit{\Gamma}_2, \mathit{\Delta}_2, x{:}B) \end{array}$$

and thus

$$\begin{array}{rcl} u & \leq & \max(0, \mathit{\Phi}_{\mathcal{V},\mathcal{H}}(\mathit{\Gamma}_1, \mathit{\Delta}_1) + \mathit{\Phi}_{\mathcal{V}',\mathcal{H}_1}(\mathit{\Gamma}_2, \mathit{\Delta}_2) + s_1 + K^{\text{let}}_1) \\ & \overset{(\mathit{Lem.}\ 4)}{\leq} & \max(0, \mathit{\Phi}_{\mathcal{V},\mathcal{H}}(\mathit{\Gamma}_1, \mathit{\Delta}_1) + \mathit{\Phi}_{\mathcal{V},\mathcal{H}}(\mathit{\Gamma}_2, \mathit{\Delta}_2) + s_1 + K^{\text{let}}_1) \\ & \overset{(43,44)}{\leq} & \mathit{\Phi}_{\mathcal{V},\mathcal{H}}(\mathit{\Gamma}) + q \end{array}$$

Finally, it follows from Prop. 1 applied to (47), (49), and (42) that $u \geq p$.