Resource-Guided Program Synthesis

Tristan Knoth  
University of California, San Diego  
tknoth@ucsd.edu

Nadia Polikarpova  
University of California, San Diego  
npolikarpova@ucsd.edu

Di Wang  
Carnegie Mellon University  
diw3@cs.cmu.edu

Jan Hoffmann  
Carnegie Mellon University  
jhoffmann@cmu.edu

Abstract
This article presents resource-guided synthesis, a technique for synthesizing recursive programs that satisfy both a functional specification and a symbolic resource bound. The technique is type-directed and rests upon a novel type system that combines polymorphic refinement types with potential notations of automatic amortized resource analysis. The type system enables efficient constraint-based type checking and can express precise refinement-based resource bounds. The proof of type soundness shows that synthesized programs are correct by construction. By tightly integrating program exploration and type checking, the synthesizer can leverage the user-provided resource bound to guide the search, eagerly rejecting incomplete programs that consume too many resources. An implementation in the resource-guided synthesizer ReSyn is used to evaluate the technique on a range of recursive data structure manipulations. The experiments show that ReSyn synthesizes programs that are asymptotically more efficient than those generated by a resource-agnostic synthesizer. Moreover, synthesis with ReSyn is faster than a naive combination of synthesis and resource analysis. ReSyn is also able to generate implementations that have a constant resource consumption for fixed input sizes, which can be used to mitigate side-channel attacks.

CCS Concepts  
• Software and its engineering → Automatic programming;  
• Theory of computation → Automated reasoning;

Keywords  
Program Synthesis, Automated Amortized Resource Analysis, Refinement Types

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ACM Reference Format:

1 Introduction
In recent years, program synthesis has emerged as a promising technique for automating low-level aspects of programming [24, 61, 66]. Synthesis technology enables users to create programs by describing desired behavior with input-output examples [18–20, 22, 47, 60, 70, 71], natural language [72], and partial or complete formal specifications [35, 39, 51, 52, 63]. If the input is a formal specification, synthesis algorithms can not only create a program but also a proof that the program meets the given specification [39, 51, 52, 63].

One of the greatest challenges in software development is to write programs that are not only correct but also efficient with respect to memory usage, execution time, or domain specific resource metrics. For this reason, automatically optimizing program performance has long been a goal of synthesis, and several existing techniques tackle this problem for low-level straight-line code [9, 49, 50, 56, 57] or add efficient synchronization to concurrent programs [11, 12, 21, 28]. However, the developed techniques are not applicable to recent advances in the synthesis of high-level looping or recursive programs manipulating custom data structures [22, 35, 39, 47, 51, 52]. These techniques lack the means to analyze and understand the resource usage of the synthesized programs. Consequently, they cannot take into account the program’s efficiency and simply return the first program that arises during the search and satisfies the functional specification.

In this work, we study the problem of synthesizing high-level recursive programs given both a functional specification of a program and a bound on its resource usage. A naive solution would be to first generate a program using conventional program synthesis and then use existing automatic static resource analyses [15, 32, 48] to check whether its resource usage satisfies the bound. Note, however, that for recursive programs, both synthesis and resource analysis are undecidable in theory and expensive in practice. Instead, in this paper we propose resource-guided synthesis: an approach that tightly
integrates program synthesis and resource analysis, and uses the resource bound to guide the synthesis process, generating programs that are efficient by construction.

**Type-Driven Synthesis** In a nutshell, the idea of this work is to combine type-driven program synthesis, pioneered in the work on Synquid [51], with type-based automatic amortized resource analysis (AARA) [31, 33, 34, 37] as implemented in Resource Aware ML (RaML) [30]. Type-driven synthesis and AARA are a perfect match because they are both based on decidable, constraint-based type systems that can be easily checked with off-the-shelf constraint solvers.

In Synquid, program specifications are written as refinement types [41, 68]. The key to efficient synthesis is round-trip type checking, which uses an SMT solver to aggressively prune the search space by rejecting partial programs that do not meet the specification (see Sec. 2.1). Until now, types have only been used in the context of synthesis to specify functional properties.

AARA is a type-based technique for automatically deriving symbolic resource bounds for functional programs. The idea is to add resource annotations to data types, in order to specify a potential function that maps values of that type to non-negative numbers. The type system ensures that the initial potential is sufficient to cover the cost of the evaluation. By a priori fixing the shape of the potential functions, type inference can be reduced to linear programming (see Sec. 2.2).

**The Re² Type System** The first contribution of this paper is a new type system, which we dub Re²—for refinements and resources—that combines polymorphic refinement types with AARA (Sec. 3). Re² is a conservative extension of Synquid’s refinement type system and RaML’s affine type system with linear potential annotations. As a result, Re² can express logical assertions that are required for effectively specifying program synthesis problems. In addition, the type system features annotations of numeric sort in the same refinement language to express potential functions. Using such annotations, programmers can express precise resource bounds that go beyond the template potential functions of RaML.

The features that distinguish Re² from other refinement-based type systems for resource analysis [15, 48, 53] are (1) the combination of logical and quantitative refinements and (2) the use of AARA, which simplifies resource constraints and naturally applies to non-monotone resources like memory that can become available during the execution. These features also pose nontrivial technical challenges: the interaction between substructural and dependent types is known to be tricky [42, 43], while polymorphism and higher-order functions are challenging for AARA (one solution is proposed in [37], but their treatment of polymorphism is not fully formalized).

In addition to the design of Re², we prove the soundness of the type system with respect to a small-step cost semantics. In the formal development, we focus on a simple call-by-value functional language with Booleans and lists, where type refinements are restricted to linear inequalities over lengths of lists. However, we structure the formal development to emphasize that Re² can be extended with user-defined data types, more expressive refinements, or non-linear potential annotations. The proof strategy itself is a contribution of this paper. The type soundness of the logical refinement part of the system is inspired by TiML [48]. The main novelty is the soundness proof of the potential annotations using a small-step cost semantics instead of RaML’s big-step evaluation semantics.

**Type-Driven Synthesis with Re²** The second contribution of this paper is a resource-guided synthesis algorithm based on Re². In Sec. 4, we first develop a system of synthesis rules that prescribe how to derive well-typed programs from Re² types, and prove its soundness wrt. the Re² type system. We then show how to algorithmically derive programs using a combination of backtracking search and constraint solving. In particular this requires solving a new form of constraints we call resource constraints, which are constrained linear inequalities over unknown numeric refinement terms. To solve resource constraints, we develop a custom solver based on counter-example guided inductive synthesis [62] and SMT [17].

**The ReSyn Synthesizer** The third contribution of this paper is the implementation and experimental evaluation of the first resource-aware synthesizer for recursive programs. We implemented our synthesis algorithm in a tool called ReSyn, which takes as input (1) a goal type that specifies the logical refinements and resource requirements of the program, and (2) types of components (i.e. library functions that the program may call). ReSyn then synthesizes a program that provably meets the specification (assuming the soundness of components).

To evaluate the scalability of the synthesis algorithm and the quality of the synthesized programs, we compare ReSyn with baseline Synquid on a variety of general-purpose data structure operations, such as eliminating duplicates from a list or computing common elements between two lists. The evaluation (Sec. 5) shows that ReSyn is able to synthesize programs that are asymptotically more efficient than those generated by Synquid. Moreover, the tool scales better than a naive combination of synthesis and resource analysis.

## 2 Background and Overview

This section provides the necessary background on type-driven program synthesis (Sec. 2.1) and automatic resource analysis (Sec. 2.2). We then describe and motivate their combination in Re² and showcase novel features of the type system (Sec. 2.3). Finally, we demonstrate how Re² can be used for resource-guided synthesis (Sec. 2.4).

### 2.1 Type-Driven Program Synthesis

Type-driven program synthesis [51] is a technique for automatically generating functional programs from their high-level specifications expressed as refinement types [41, 54].
Resource-Guided Program Synthesis

```plaintext
1  common = \lambda l1 . \lambda l2 . match l1 with Nil → Nil
2  Cons x xs → if ¬(member x l2)
3    then common xs l2
4  else Cons x (common xs l2)
```

Figure 1. Synthesized program that computes common elements between two lists

For example, a programmer might describe a function that computes the common elements between two lists using the following type signature:

```
common :: List a → List l1 → List l2 → \{\nu : List a | \nu = \text{elems } l1 \cap \text{elems } l2\}
```

Here, the return type of `common` is refined with the predicate \(\text{elems } \nu = \text{elems } l1 \cap \text{elems } l2\), which restricts the set of elements of the output list \(\nu\) to be the intersection of the sets of elements of the two arguments. Here \(\text{elems}\) is a user-defined logic-level function, also called `measure` [38, 68]. In addition to the `synthesis goal` above, the synthesizer takes as input a `component library`: signatures of data constructors and functions it can use. In our example, the library includes the list constructors `Nil` and `Cons` and the function `member :: x : a → l : List a → \{\text{Bool} | \nu = \{x \in \text{elems } l\}\}` which determines whether a given value is in the list. Given this goal and components, the type-driven synthesizer `SYNQUID` [51] produces an implementation of `common` in Fig. 1.

### The Synthesis Mechanism

Type-driven synthesis works by systematically exploring the space of programs that can be built from the component library and validating candidate programs against the goal type using a variant of liquid type inference [54]. To validate a program against a refinement type, liquid type inference generates a system of `subtyping constraints` over refinement types. The subtyping constraints are then reduced to implications between refinement predicates. For example, checking `common xs l2` in line 3 of Fig. 1 against the goal type reduces to validating the following implication:

\[
\text{(elems } l1 = \{x\} \cup \text{elems } xs\} \land \neg (x \in \text{elems } l2) \land \text{elems } v = \text{elems } xs \cap \text{elems } l2 \implies \text{elems } v = \text{elems } l1 \cap \text{elems } l2
\]

Since this formula belongs to a decidable theory of uninterpreted functions and arrays, its validity can be checked by an SMT solver [17]. In general, the generated implications may contain unknown predicates. In this case, type inference reduces to a system of `constrained horn clauses` [6], which can be solved via predicate abstraction.

### Synthesis and Program Efficiency

The program in Fig. 1 is correct, but not particularly efficient: it runs roughly in time \(n \cdot m\), where \(m\) is the length of \(l1\) and \(n\) is the length of \(l2\), since it calls the `member` function (a linear scan) for every element of \(l1\). The programmer might realize that keeping the input lists

---

Footnotes:

1. Hereafter the bound variable of the refinement is always called \(\nu\) and the binding is omitted.
2. Following `SYNQUID`, our language imposes an implicit constraint on all type variables to support equality and ordering. Hence, they cannot be instantiated with arrow types. This could be lifted by adding type classes.
3. In this section we assume for simplicity that the resource of interest is the number of recursive calls. Both AARA and our type system support user-defined cost metrics (see Sec. 3 for details).
Potential Annotations  
AARA is inspired by the potential method for manually analyzing the worst-case cost of a sequence of operations [64]. It uses annotated types to introduce potential functions that map program states to non-negative numbers. To derive a bound, we have to statically ensure that the potential at every program state is sufficient to cover the cost of the next transition and the potential of the following state. In this way, we ensure that the initial potential is an upper bound on the total cost.

The key to making this approach effective is to closely integrate the potential functions with data structures [34, 37]. For instance, in RaML the type $L^1(\text{int})$ stands for a list that contains one unit of potential for every element. This type defines the potential function $\phi(l: L^1(\text{int})) = 1 - |l|$. The potential can be used to pay for a recursive call (or, in general, cover resource usage) or to assign potential to other data structures.

Bound Inference  
Potential annotations can be derived automatically by starting with a symbolic type derivation that contains fresh variables for the potential annotations of each type, and applying syntax directed type rules that impose local constraints on the annotations. The integration of data structures and potential ensures that these constraints are linear even for polynomial potential annotations.

2.3 Bounding Resources with Re$^2$

To reason about resource usage in type-driven synthesis, we integrate AARA’s potential annotations and refinement types into a novel type system that we call Re$^2$. In Re$^2$, a refinement type can be annotated with a potential term $\phi$ of numeric sort, which is drawn from the same logic as refinements. Intuitively, the type $R^\phi$ denotes values of refinement type $R$ with $\phi$ units of potential. In the rest of this section we illustrate features of Re$^2$ on a series of examples, and delay formal treatment to Sec. 3.

With potential annotations, users can specify that common

\begin{verbatim}
common':: l1: SList a1 \rightarrow l2: SList a1
\rightarrow (\forall:\text{List }a | \text{elems } v = \text{selems } l1 \land \text{selems } l2)
\end{verbatim}

This type assigns one unit of potential to every element of the arguments $l1$ and $l2$, and hence only allows making one recursive call per element of each list. Whenever resource annotations are omitted, the potential is implicitly zero: for example, the elements of the result carry no potential.

Our type checker uses the following reasoning to argue that this potential is sufficient to cover the efficient implementation in Fig. 2. Consider the recursive call in line 4, which has a cost of one. Pattern-matching $l1$ against SCons $\times xs$ transfers the potential from $l1$ to the binders, resulting in types $x : a1$ and $xs : SList (\{a | x < v\})$. The unit of potential associated with $x$ can now be used to pay for the recursive call. Moreover, the types of the arguments, $xs$ and $l2$, match the required type $SList a1$, which guarantees that the potential stored in the tail

append::xs:List a1 \rightarrow ys:List a
\rightarrow (\text{List } a | \text{len } v = \text{len } xs + \text{len } ys)

triple::l:List Int^2 \rightarrow (\{\text{List } n | \text{len } v = 3*(\text{len } l)\})
triple = \lambda l. \text{append } (\text{append } l l)

tripleSlow::l:List Int^3 \rightarrow (\{\text{List } n | \text{len } v = 3*(\text{len } l)\})
tripleSlow = \lambda l. \text{append } (\text{append } l l) l

Figure 3. Append three copies of a list. The type of append specifies that it returns a list whose length is the sum of the lengths of its arguments. It also requires one unit of potential on each element of the first list. Moreover, append has a polymorphic type and can be applied to lists with different element types, which is crucial for type-checking tripleSlow.

and the second list are sufficient to cover the rest of the evaluation. Other recursive calls are checked in a similar manner.

Importantly, the inefficient implementation in Fig. 1 would not type-check against this signature. Assuming that member is soundly annotated with

\begin{verbatim}
member::x:a \rightarrow l:List a \rightarrow (\text{Bool } v = (x \text{ in } \text{elems } l))
\end{verbatim}

requiring a unit of potential per element of $l$), the guard in line 2 consumes all the potential stored in $l2$; hence the occurrence of $l2$ in line 3 has the type $List a2$, which is not a subtype of $List a1$.

Dependent Potential Annotations  
In combination with logical refinements and parametric polymorphism, this simple extension to the Synquid’s type system turns out to be surprisingly powerful. Unlike in RaML, potential annotations in Re$^2$ can be dependent, i.e. mention program variables and the special variable $v$. Dependent annotations can encode fine-grained bounds, which are out of reach for RaML. As one example, consider function range $\text{a \ b}$ that builds a list of all integers between $a$ and $b$; we can express that it takes at most $b - a$ steps by giving the argument $b$ a type $\{\text{Int } | v \geq a\}^{b-a}$. As another example, consider insertion into a sorted list insert $\times xs$; we can express that it takes at most as many steps as there are elements in $xs$ that are smaller than $x$, by giving $xs$ the type $SList a^{\text{le}(v < x, 1.0)}$ (i.e. only assigning potential to elements that are smaller than $x$). These fine-grained bounds are checked completely automatically in our system, by reduction to constraints in SMT-decidable theories.

Polymorphism  
Another source of expressiveness in Re$^2$ is parametric polymorphism: since potential annotations are attached to types, type polymorphism gives us resource polymorphism for free. Consider two functions in Fig. 3, triple and tripleSlow, which implement two different ways to append a list $l$ to two copies of itself. Both of them make use of a component function append, whose type indicates that it makes a linear traversal of its first argument. Intuitively,
triple is more efficient that tripleSlow because in the former both calls to append traverse a list of length \( n \), whereas in the latter the outer call traverses a list of length \( 2n \). This difference is reflected in the signatures of the two functions: tripleSlow requires three units of potential per list element, while triple only requires two.

Checking that tripleSlow satisfies this bound is somewhat nontrivial because the two applications of append must have different types: the outer application must return \( \text{List } \text{Int} \), while the inner application must return \( \text{List } \text{Int}^1 \) (i.e. carry enough potential to be traversed by append). RaML’s monomorphic type system is unable to assign a single general type to append, which can be used at both call sites. So the function has be reanalyzed at every (monomorphic) call site. ReSyn, on the other hand, handles this example out of the box, since the type variable \( a \) in the type of append can be instantiated with \( \text{Int} \) for the outer occurrence and with \( \text{Int}^1 \) for the inner occurrence, yielding the type

\[
\text{xs}: \text{List } \text{Int}^2 \rightarrow \text{ys}: \text{List } \text{Int}^1 \rightarrow (\text{List } \text{Int}^1 | ...)
\]

As a final example, consider the standard map function:

\[
\text{map} :: (a \rightarrow b) \rightarrow \text{List } a \rightarrow \text{List } b
\]

Although this type has no potential annotations, it implicitly tells us something about the resource behavior of map: namely, that map applies a function to each list element at most once. This is because a can be instantiated with a type with an arbitrary amount of potential, and the only way to pay for this potential is with a list element (which also has type \( a \)).

### 2.4 Resource-guided Synthesis with ReSyn

We have extended SYNQUID with support for ReSyn types in a new program synthesizer ReSyn. Given a resource-annotated signature for common' from Sec. 2.3 and a component library that includes member, ReSyn is able to synthesize the efficient implementation in Fig. 2. The key to efficient synthesis is type-checking each program candidate incrementally as it is being constructed, and discarding an ill-typed program prefix as early as possible. For example, while enumerating candidates for the function common', we can safely discard the inefficient version from Fig. 1 even before constructing the second branch of the conditional (because the first branch together with the guard use up too many resources). Hence, as we explain in more detail in Sec. 4, a key technical challenge in ReSyn has been a tight integration of resources into SYNQUID’s round-trip type checking mechanism, which aggressively propagates type information top-down from the goal and solves constraints incrementally as they arise.

#### Termination Checking

In addition to making the synthesizer resource-aware, ReSyn types also subsume and generalize SYNQUID’s termination checking mechanism. To avoid generating diverging functions, SYNQUID uses a simple termination metric (the tuple of function’s arguments), and checks that this metric decreases at every recursive call. Using this metric,

\[
\begin{align*}
  a & := x | \text{true} | \text{false} | \text{nil} | \text{cons}(\hat{a}_n,a_i) \\
  \hat{a} & := a | \lambda (x.e_0) | \text{fix}(f.x.e_0) \\
  e & := \hat{a} | \text{if}(\hat{a}_0,e_1,e_2) | \text{matl}(a_0,e_1,x_0,x_1.e_2) | \text{app}(\hat{a}_1,\hat{a}_2) \\
  & \quad | \text{let}(e_1,x.e_2) | \text{impossible}(\text{tick}(c,e_0)) \\
  v & := \text{true} | \text{false} | \text{nil} | \text{cons}(v_h,v_i) | \lambda (x.e_0) | \text{fix}(f.x.e_0)
\end{align*}
\]

**Figure 4.** Syntax of the core calculus

SYNQUID is not able to synthesize the function range from Sec. 2.3, because it requires a recursive call that decreases the difference between the arguments, \( b - a \). In contrast, ReSyn need not reason explicitly about termination, since potential annotations already encode an upper bound on the number of recursive calls. Moreover, the flexibility of these annotations enables ReSyn to synthesize programs that require nontrivial termination metrics, such as range.

### 3 The Re² Type System

In this section, we define a subset of Re² as a formal calculus to prove type soundness. This subset includes Booleans that are refined by their values, and lists that are refined by their lengths. The programs in Sec. 1 and Sec. 2 use SYNQUID’s surface syntax. The gap from the surface language to the core calculus involves inductive types and refinement-level measures. The restriction to this subset in the technical development is only for brevity and proofs carry over to all the features of SYNQUID.

**Syntax**Fig. 4 presents the grammar of terms in Re² via abstract binding trees [29]. The core language is basically the standard lambda calculus augmented with Booleans and lists. A value \( v \in \text{Val} \) is either a boolean constant, a list of values, or a function. Expressions in Re² are in a-normal-form [55], which means that syntactic forms occurring in non-tail position allow only atoms \( a \in \text{Atom} \), i.e., variables and values; this restriction simplifies typing rules for applications, as we explain below. We identify a subset SimpAtom of Atom that contains atoms interpretable in the refinement logic. Intuitively, the value of an \( a \in \text{SimpAtom} \) should be either a Boolean or a list. The syntactic form impossible is introduced as a placeholder for unreachable code, e.g., the else-branch of a conditional whose predicate is always true.

The syntactic form \( \text{tick}(c,e_0) \) is used to specify resource usage, and it is intended to cost \( c \in \mathbb{Z} \) units of resource and then reduce to \( e_0 \). If the cost \( c \) is negative, then \(-c\) units of resource will become available in the system. tick terms support flexible user-defined cost metrics: for example, to count recursive calls, the programmer may wrap every such call in \( \text{tick}(\cdot,\cdot) \); to keep track of memory consumption, they might wrap every data constructor in \( \text{tick}(c,\cdot) \), where \( c \) is the amount of memory that constructor allocates.

**Operational Semantics** The resource usage of a program is determined by a small-step operational cost semantics. The semantics is a standard one augmented with a resource
The *multi-step* evaluation relation $\Rightarrow^*$ is the reflexive transitive closure of $\Rightarrow$. The judgment $(e, q) \Rightarrow^* (e', q')$ expresses that with $q$ units of available resources, $e$ evaluates to $e'$ without running out of resources and $q'$ resources are left. Intuitively, the high-water mark resource usage of an evaluation $e$ is the minimal $q$ such that $(e, q) \Rightarrow^* (e', q')$. For monotone resources like time, the cost is the sum of costs of all the evaluated tick expressions. In general, this net cost is invariant, that is, $p - p' = q - q'$ if $(e, p) \Rightarrow^* (e', p')$ and $(e, q) \Rightarrow^* (e', q')$, where $\Rightarrow^*$ is the relation obtained by self-composing $\Rightarrow$ for $n$ times.

**Refinements** We now combine SYNQUI's type system with AARA to reason about resource usage. Fig. 5 shows the syntax of the Re² type system. Refinements $\psi$ are distinct from program terms and classified by sorts $\Delta$. Re² sorts include Booleans $\mathbb{B}$, natural numbers $\mathbb{N}$, and *uninterpreted symbols* $\delta_a$. Refinements can be logical formulas and linear expressions, which may reference program variables. Logical refinements $\psi$ have sort $\mathbb{B}$, while potential annotations $\phi$ have sort $\mathbb{N}$. Re² interprets a variable of Boolean type as its value, list type as its length, and type variable $\alpha$ as an uninterpreted symbol with a corresponding sort $\delta_a$. We use the following interpretation $I(\cdot)$ to reflect interpretable atoms $a \in \text{SimpAtom}$ in the refinement logic:

$$
I(x) = x \\
I(\text{true}) = \top \\
I(\text{false}) = \bot \\
I(\text{cons}(\_a)) = I(a_i) + 1
$$

**Types** We classify types into four categories. Base types $B$ include Booleans, lists and type variables. Type variables $\alpha$ are annotated with a *multiplicity* $m \in \mathbb{Z}_0^+ \cup \{\infty\}$, which denotes an upper bound on the number of usages of a variable like in bounded linear logic [23]. For example, $L(2 \cdot \alpha)$ denotes a universal list whose elements can be used at most twice.

Refinement types are *subset types* and *dependent arrow types*. The inhabitants of the subset type $\{B \mid \psi\}$ are values of type $B$ that satisfy the refinement $\psi$. The refinement $\psi$ is a logical predicate over program variables and a special *value variable* $\nu$, which does not appear in the program and stands for the inhabitant itself. For example, $\{\text{bool} \mid \nu\}$ is a type of true, and $\{L(\text{bool}) \mid \nu \leq 5\}$ represents Boolean lists of length at most 5. Dependent arrow types $x:T_x \rightarrow T$ are function types whose return type may reference the formal argument $x$. As type variables, these function types are also annotated with a multiplicity $m \in \mathbb{Z}_0^+ \cup \{\infty\}$ restricting the number of times the function may be applied.

To apply the potential method of amortized analysis [65], we need to define potentials with respect to the data structures in the program. We introduce *resource-annotated types* as a refinement type augmented with a potential annotation, written $R^{\phi}$. Intuitively, $R^{\phi}$ assigns $\phi$ units of potential to values of the refinement type $R$. The potential annotation $\phi$ may also reference the value variable $\nu$. For example, $L(\text{bool})^{3 \times \phi}$ describes Boolean lists $\ell$ with $5|\ell|$ units of potential where $|\ell|$ is the length of $\ell$. The same potential can be expressed by assigning 5 units of potential to every element using the type $L(\text{bool})^5$.

Type schemas represent (possibly) polymorphic types. Note that the type quantifier $\forall$ can only appear outermost in a type.

Similar to SYNQUI, we introduce a notion of *scalar types*, which are resource-annotated base types refined by logical constraints. Intuitively, interpretable atoms are scalars and Re² only allows the refinement-level logic to reason about values of scalar types. We will abbreviate $\cdot \cdot \cdot \alpha \cdot \cdot \cdot$ as $\cdot \cdot \cdot$, $\{B \mid \top\}$ as $\infty \cdot (x:T_x \rightarrow T)$ as $x:T_x \rightarrow T$, and $R^{\phi}$ as $R$.

**Typing Rules** In Re², the *typing context* $\Gamma$ is a sequence of variable bindings $x:S$, type variables $\alpha$, path conditions $\psi$, and free potentials $\phi$. Our type system consists of five judgments: sorting, well-formedness, subtyping, sharing, and typing. We omit sorting and well-formedness rules and include them in the technical report [40]. The sorting judgment $\Gamma \vdash \psi \in \Delta$ states that a refinement $\psi$ has a sort $\Delta$ under a context $\Gamma$. A type $S$ is said to be well-formed under a context $\Gamma$, written $\Gamma \vdash S$, if every referenced variable in it is in the correct scope.

Fig. 6 presents selected typing rules for Re². The typing judgment $\Gamma \vdash e : S$ states that the expression $e$ has type $S$ in context $\Gamma$. The intuitive meaning is that if there is at least the amount resources as indicated by the potential in the context $\Gamma$ then this suffices to evaluate $e$ to a value $\nu$, and after the evaluation there are at least as many resources available as indicated by the potential in $S$. The auxiliary typing judgment $\Gamma \vdash a : B$ assigns base types to interpretable atoms. Atomic typing is useful in the rule (T-SimpAtom), which uses the interpretation $I(\cdot)$ to derive a most precise refinement type for interpretable atoms.

The *subtyping* judgment $\Gamma \vdash T_1 :< \cdot T_2$ is defined in a standard way, with the extra requirement that the potential in $T_1$ should be greater than or equal to that in $T_2$. Subtyping is often used to ”forget” some program variables in the type to ensure the result type does not reference any locally introduced variable.
The ability to derive precise types for dependently typed applications motivates the use of a normal-form in Re². Figure 6. Selected typing rules of the Re² type system

e.g., the result type of let(e₁, x, e₂) cannot have x in it and the result type of mat(a₀, e₁, e₂) cannot reference e₀ or e₁.

To reason about logical invariants, we introduce validity checking, written Γ ⊨ ψ, to state that a logical refinement ψ is always true under any instance of the context Γ. The validity checking relation is established upon a denotational semantics for refinements. Validity checking in Re² is decidable because it can be reduced to Presburger arithmetic. The full development of validity checking is included in the technical report [40].

We reason about inductive invariants for lists in rule (T-MatL), using interpretation ℐ(·). In our formalization, lists are refined by their length thus the invariants are: (i) nil has length 0, and (ii) the length of cons(a_i) is the length of a_i plus one. The type system can be easily enriched with more refinements and data types (e.g., the elements of a list are the union of its head and those of its tail) by updating the interpretation ℐ(·) as well as the premises of rule (T-MatL).

Finally, notable are the two typing rules for applications: (T-App) and (T-App-SimpAtom). In the former case, the function return type T does not mention x, and hence can be directly used as the type of the application (this is the case e.g. for all higher-order applications, since our well-formedness rules prevent functions from appearing in refinements). In the latter case, T mentions x, but luckily any argument of a scalar type must be a simple atom a, so we can substitute x with its interpretation ℐ(a). The ability to derive precise types for dependent applications motivates the use of a normal-form in Re².
**Resources** The rule (T-CONSUME-P) states that an expression tick(c,e₀) is only well-typed in a context that contains a free potential term c. To transform the context into this form, we can use the rule (S-TRANSFER) to transfer potential within the context between type variables and free potential terms, as long as we can prove that the total amount of potential remains the same. For example, the combination of (S-TRANSFER) and (S-RELAX) allows us to derive both x : bool₁ + x : bool and x : bool₁ + tick(1,x) : bool (but not x : bool₁ + tick(1,x) : bool₁).

The typing rules of Re² form an affine type system [69]. To use a program variable multiple times, we have to introduce explicit sharing to ensure that the program cannot gain potential. The sharing judgment Γ ⊢ S \ Y S₁ | S₂ means that in the context Γ, the potential indicated by S is apportioned into two parts to be associated with S₁ and S₂. We extend this notion to context sharing, written Γ ⊢ Γ ′ | Γ₂, which states that Γ₁,Γ₂ has the same sequence of bindings as Γ, but the potentials of type bindings in Γ are shared point-wise, and the free potentials in the Γ are also split. A special context sharing Γ ⊢ ∏ Γ ′ | Γ is used in the typing rules (T-Relax) and (T-Fix) for functions. The self-sharing indicates that the function can only reference potential-free free variables in the context. This is also used to ensure that the program cannot gain more potential through free variables by applying the same function multiple times.

Restricting functions to be defined under potential-free contexts is undesirable in some situations. For example, a curried function of type x : Tₓ → y : Tᵧ → T might require nonzero units of potential on its first argument x, which is not allowed by rule (T-Relax) or (T-Fix) on the inner function type y : Tᵧ → T. We introduce another rule (T-Relax-Lin) to relax the restriction. The rule associates a multiplicity m with the function type, which denotes the number of times that the function could be applied. Instead of context self-sharing, we require the potential in the context to be enough for m function applications. Note that in ReSyn’s surface syntax used in the Sec. 2, every curried function type implicitly has multiplicity 1 on the inner function: x : Tₓ → 1 \ y : Tᵧ → T).

**Example** Recall the function `triple` from Fig. 3, which can be written as follows in Re² core syntax:

\[
\begin{align*}
\text{triple} & \equiv \ell : L(\text{bool}^2) \rightarrow L(\text{bool}) | v = 3 \times \ell \\
\text{triple} & \equiv \lambda (\ell, (\text{let}(\text{app(append}, \ell), \ell')).
\end{align*}
\]

Next, we illustrate how Re² uses the signature of append:

\[
\text{append} \equiv \forall \alpha. \forall \nu. (\ell : L(\text{bool}) | v = \nu) \rightarrow (L(\text{bool}) | v = \nu)
\]

to justify the resource bound 2|ℓ| on `triple`. Suppose Γ is a typing context that contains the signature of append. The argument ℓ is used three times, so we need to use sharing relations to apportion the potential of ℓ. We have Γ ⊢ L(\text{bool}_1^2) \ Y L(\text{bool}_1) \ Y L(\text{bool}_1) \ X L(\text{bool}_1) \ Y L(\text{bool}_1), and we assign L(\text{bool}_1), L(\text{bool}_1), and L(\text{bool}_1) to the three occurrences of ℓ respectively in the order they appear in the program. To reason about e₁ = app(append, ℓ), ℓ, we instantiate append with α : bool, inferring its type as

\[
xs : L(\text{bool}_1) \rightarrow 1 \cdot (ys : L(\text{bool}) \rightarrow (L(\text{bool}) | v = xs + ys))
\]

and by (T-APP-SIMPATOM) we derive the following:

\[
Γ, ℓ : L(\text{bool}_1), ℓ' : T₁ + e₁ : L(\text{bool}_1) | v = ℓ + ℓ' \frac{Γ, ℓ}{}.
\]

We then can typecheck e₂ = app(append, ℓ), ℓ’ with the same instantiation of append:

\[
Γ, ℓ : L(\text{bool}_1), ℓ' : T₁ + e₂ : L(\text{bool}_1) | v = e₁ + (xs + (xs + xs)) \frac{Γ, ℓ}{}.
\]

(Where T₁ is the type of e₁). Finally, by subtyping and the following valid judgment in the refinement logic

\[
Γ, ℓ : L(\text{bool}_1), v : L(\text{bool}) \rightarrow ℓ = ℓ + ℓ' \frac{Γ, ℓ}{} \rightarrow v = 3 \times ℓ,
\]

we conclude Γ ⊢ triple : L(\text{bool}_1) \rightarrow (L(\text{bool}) | v = 3 \times ℓ).

**Soundness** The type soundness for Re² is based on progress and preservation. The progress theorem states that if we derive a bound q for an expression e with the type system and p ≥ q resources are available, then (e,p) can make a step if e is not a value. In this way, progress shows that resource bounds are indeed bounds on the high-water mark of the resource usage since states (e,p) in the small step semantics can be stuck based on resource usage if, for instance, p = 0 and e = tick(1,e’).

**Theorem 1** (Progress). If q + e : S and p ≥ q, then either e ∈ Val or there exist e’ and p’ such that (e,p) → (e’,p’).

**Proof.** By strengthening the assumption to Γ ⊢ e : S where Γ is a sequence of type variables and free potentials, and then induction on Γ ⊢ e : S.

The preservation theorem accounts for resource consumption by relating the left over resources after a computation to the type judgment of the new term.

**Theorem 2** (Preservation). If q + e : S, p ≥ q and (e,p) → (e’,p’), then p’ ⊢ e’ : S.

**Proof.** By strengthening the assumption to Γ ⊢ e : S where Γ is a sequence of free potentials, and then induction on Γ ⊢ e : S, followed by inversion on the evaluation judgment (e,p) → (e’,p’).

The proof of preservation makes use of the following crucial substitution lemma.

**Lemma 1** (Substitution). If Γ₁, x : B | ψ₁, Γ’ ⊢ e : S, Γ₂ ⊢ t : B | ψ₂, t ∈ Val and Γ ⊢ Γ₁ | Γ₂, then Γ, [Γ₂][t/x][ψ₁]Γ’ ⊢ [Γ₁][t/x]e : [Γ₂][t/x]S.

**Proof.** By induction on Γ₁, x : B | ψ₁, Γ’ ⊢ e : S.

Since we found the purely syntactic soundness statement about results of computations (they are well-typed values) somewhat unsatisfactory, we also introduced a denotational notion of consistency. For example, a list of values ℓ = [v₁,...,vₜ]is consistent with q + ℓ : L({\text{bool}|v}) | v = q, if q ≥ n + 5 and each value vᵢ of the list is false. We then show that well-typed values are consistent with their typing judgement.
where

\[
D ::= \cdot \mid D, x \leftarrow e
\]

\[
\hat{e} ::= e \mid \mathtt{app}(x, \bar{o}) \mid \mathtt{matl}(x, \bar{o}, x_h, x_l, \bar{c}) \mid \mathtt{lets}(D, \hat{e})
\]

\[
T ::= R^\emptyset \mid ?
\]

Figure 7. Extended syntax

Lemma 2 (Consistency). If \(q \vdash v : S\), then \(v\) satisfies the conditions indicated by \(S\) and \(q\) is greater than or equal to the potential stored in \(v\) with respect to \(S\).

As a result, we derive the following theorem.

Theorem 3 (Soundness). If \(q \vdash e : S\) and \(p \geq q\) the either

- \((e, p) \rightarrow \langle v, p' \rangle\) and \(v\) is consistent with \(p' \vdash v : S\) or
- for every \(n\) there is \(\langle e', p' \rangle\) such that \((e, p) \rightarrow^n \langle e', p' \rangle\).

Complete proofs can be found in the technical report [40].

Inductive Datatypes and Measures We can generalize our development of list types for inductive types \(\mu X.C : T \times X^k\), where \(C\) is the constructor name, \(T\) is the element type that does not contain \(X\), and \(X^k\) is the \(k\)-element product type \(X \times X \times \cdots \times X\). The introduction rules and elimination rules are almost the same as (T-Nil), (T-Cons) and (T-MatL), respectively, except that we need to capture inductive invariants for each constructor \(C\) in the rules correspondingly. In SYNQID, these invariants are specified by inductive measures that map values to refinements. We can introduce new sorting rules for inductive types to embed values as their related measures in the refinement logic.

Constant Resource Our type system infers upper bounds on resource usage. Recently, AARA has been generalized to verify constant-resource behavior [46]. A program is said to be constant-resource if its executions on inputs of the same size consume the same amount of resource. We can adapt the technique in [46] to \(\text{Re}^2\) by (i) changing the subtyping rules to keep potentials invariant (i.e. replacing \(\geq\) with \(=\) in (Sub-TVAr), (Sub-Arrow), (Sub-Pot)), and (ii) changing the rule (Simp-Atom-Var) to require \(\phi = 0\). Based on the modified type system, our synthesis algorithm can also synthesize constant-time implementations (see Sec. 5.2 for more details).

4 Type-Driven Synthesis with \(\text{Re}^2\)

In this section, we first show how to turn the type checking rules of \(\text{Re}^2\) into synthesis rules, and then leverage these rules to develop a synthesis algorithm.

4.1 Synthesis Rules

Extended Syntax To express synthesis rules, we extend \(\text{Re}^2\) with a new syntactic form \(\hat{e}\) for expression templates. As shown in Fig. 7, templates are expressions that can contain holes \(\bar{o}\) in certain positions. The flat let form \(\mathtt{lets}(D, \hat{e})\), where \(D\) is a sequence of bindings, is a shortcut for a nest of let-expressions \(\mathtt{let}(x_1, d_1, \ldots, \mathtt{let}(x_n, d_n, \hat{e}))\); we write \(\mathtt{fold}(\mathtt{lets}(D, \hat{e}))\) to convert a flat let (without holes) back to the original syntax. We also extend the language of types with an unknown type \(?\), which is used to build partially defined goal types, as explained below.

Synthesis for A-Normal-Form Our synthesis relation consists of two mutually recursive judgments: the synthesis judgment \(\Gamma \vdash \hat{e} : S \rightarrow e\) intuitively means that the template \(\hat{e}\) can be completed into an expression \(e\) such that \(\Gamma \vdash e : S\); the purpose of the auxiliary atomic synthesis judgment is explained below. Selected rules for both judgments are given in Fig. 8; the full technical development can be found in the technical report [40].

The synthesis rule (Syn-Gen) handles polymorphic goal types. The rules (Syn-Fix) and (Syn-Ans) handle arrow types and derive either a fixpoint term or an abstraction. The rule (Syn-Im) derives impossible in an inconsistent context (which may arise e.g. in a dead branch of a pattern match). The rest of the rules handle the common case when the goal type \(T\) is scalar and the context is consistent; in this case the target expression can be either a conditional, a match, or an E-term [51], i.e. a term made of variables, applications, and constructors. Special care must be taken to ensure that these expressions are in a-normal-form: generally, a-normalizing an expression requires introducing fresh variables and let-bindings for them. To retain completeness, our synthesis rules need to do the same: intuitively, in addition to an expression \(e\), a rule might also need to produce a sequence of let-bindings \(D\) that define fresh variables in \(e\). To this end, we introduce the atomic synthesis judgment \(\Gamma \vdash \hat{e} : T \rightarrow^\emptyset \mathtt{lets}(D, a)\), which synthesizes normalized E-terms, where \(a\) is an atom and each definition in \(D\) is an application or a constructor in a-normal-form.

As an example, consider the rule (Syn-Cond) for synthesizing conditionals: ideally, we would like to synthesize a guard \(e\) of type bool, and then synthesize the two branches under the assumptions that \(e\) evaluates to true and false, respectively. Recall, however, that the guard must be atomic; hence, to synthesize a well-formed conditional, we use atomic synthesis to produce a guard \(\mathtt{let}(D, x)\). Now to get a well-scoped program we must place the whole conditional inside the bindings \(D\); to that end, the second premise of (Syn-Cond) uses a nontrivial template \(\mathtt{let}(D, \mathtt{if}(x, o, c))\). The rules (Fill-Let) and (Fill-Con) handle this template by integrating it into the typing context and exposing the hole; along the way (Fill-Let) takes care of context sharing, which accounts for the potential consumed by the definitions in \(D\). Synthesis of matches works similarly using (Syn-MatL) and (Fill-MatL).

Atomic Synthesis The first four rules of atomic synthesis generate a simple atom if its type matches the goal; the rest of the rules deal with the hardest part: normalized applications. Consider the rule (ASyn-App): given a goal type \(T\) for the application \(\mathtt{app}(e_1, e_2)\), we need to construct goal types for \(e_1\) and \(e_2\), to avoid enumerating them blindly. Following SYNOID’s round-trip type checking idea, we use the type \(_? : \downarrow T\) as the goal for \(e_1\) (i.e. a function from unknown type to \(T\)). The subtyping rules for \(?\) are such that \(\Gamma \vdash (y : T_1 \rightarrow T_2) < (: _? : \rightarrow T)\)
Figure 8. Selected synthesis rules

holds if $T_2$ and $T$ agree in shape and those refinements that do not mention $y$; hence this goal type filters out those functions $e_1$ that cannot fulfill the desired goal type $T$, independently of the choice of $e_2$. One difference with SYNQUID is that the goal type for $e_1$ is linear, reflecting that we intend to use $e_1$ only once and allowing it to capture positive potential.

Similarly to the conditional case explained above, the synthesized left-hand side of the application, $e_1$, has the form $\text{lets}(D_1, x)$, and the argument $e_2$ must be synthesized inside the bindings $D_1$. These bindings are processed by (AFILL-LET), and the actual argument synthesis happens in either (AFILL-App) or (AFILL-SIMPATOM), depending on whether the argument type is a scalar. The former corresponds to a higher-order application: here $T_1$ is an arrow type, and hence the argument cannot occur in the function’s return type; in this case, synthesizing an expression of type $T_1$ must yield an abstraction or fixpoint (since $T_1$ is an arrow), both of which are atoms. The latter corresponds to a first-order application: here the return type $T'$ can mention $y$, so after synthesizing an argument of type $T_y$, we still need to check whether the resulting application $\text{lets}(D.app(x, a))$ has the right type $T$. Note how both (AFILL-App) or (AFILL-SIMPATOM) return normalized E-terms by generating a fresh variable and binding it to an application.

**Cost Metrics** In the context of synthesis we cannot rely on programmer-written tick terms to model cost. Instead in our formalization we use a simple cost metric where each function application consumes one unit of resource; hence every application generated by (AFILL-App) or (AFILL-SIMPATOM) is wrapped in tick(1, ·). Our implementation provides more flexibility and allows the programmer to annotate any arrow type with a non-negative cost $c$ to denote that applying a function of this type should incur cost $c$.

**Soundness** The synthesis rules always produce a well-typed expression (proof can be found in the technical report [40]).

Theorem 4 (Soundness of Synthesis). If $\Gamma \vdash o : S \rightarrow e$ then $\Gamma \vdash e : S$.

4.2 Synthesis Algorithm

In this section we discuss how to turn the declarative synthesis rules of Sec. 4.1 into a synthesis algorithm, which takes as input a goal type $S$, a context $\Gamma$, and a bound $k$ on the program depth, and either returns a program $e$ of depth at most $k$ such that $\Gamma \vdash e : S$, or determines that no such program exists. The core algorithm follows the recipe from prior work on type-driven synthesis [47, 51] and performs a fairly standard goal-directed backtracking proof search with $\Gamma \vdash o : e \rightarrow S$ as the top-level goal. In the rest of this section, we explain how to make such proof search feasible by reducing the core sources of non-determinism to constraint solving.
Resource-guided program synthesis

Subtyping constraints
\[
\begin{align*}
C(\Gamma \vdash m \cdot \alpha \prec \cdot m_2 \cdot \alpha) &= [\Gamma \models m_1 - m_2 \geq 0]\nC(\Gamma \vdash (B_1 \models \psi_1) < (B_2 \models \psi_1)) &= [\Gamma, \psi \models B_1 \models \psi_1 \Rightarrow B_2 \models \psi_1] \cup C(\Gamma \vdash B_1 < B_2)
C(\Gamma \vdash R_1 \cdot \psi_1 \cdot R_2 \cdot \psi_2) &= \{\Gamma, \psi \models R_1 \models \psi_1 \Rightarrow \psi_2 \geq 0\} \cup C(\Gamma \vdash R_1 < R_2)
\end{align*}
\]

Sharing constraints
\[
C(\Gamma \vdash m \cdot \gamma \succ m \cdot m_2 \cdot \alpha) = \{\Gamma \models m = -(m_1 + m_2) \geq 0
\Gamma \models m_1 + m_2 - m \geq 0\}
C(\Gamma \vdash R_1 \cdot \psi_1 \cdot R_2 \cdot \psi_2) = \{\Gamma \models \phi - (\phi_1 + \phi_2) \geq 0\} \Gamma \models \phi_1 + \phi_2 - \phi \geq 0\}
\]

Transfer constraints
\[
C(\Phi(\Gamma) = \Phi(\Gamma')) = \{\Gamma \models \Phi(\Gamma') \geq 0, \Phi(\Gamma') \models \Phi(\Gamma) \geq 0\}
\]

Figure 9. Selected cases for translating typing constraints to validity constraints.

Typing constraints The main sources of non-determinism in a synthesis derivation stem from the following premises of synthesis and typing rules: (1) whenever a given context \(\Gamma\) is shared as \(\Gamma \vdash \Gamma \subseteq \Gamma\), we need to guess how to apportion potential annotations in \(\Gamma\); (2) whenever potential in a given context \(\Gamma\) is transferred, we need to guess potential annotations in \(\Gamma'\) such that \(\Phi(\Gamma) = \Phi(\Gamma')\); and finally (3) whenever \(B \models \psi \) is used to instantiate a type variable, we need to guess both \(\phi\) and \(\psi\). All three amount to inference of unknown refinement terms of either Boolean or numeric sort. To infer these terms efficiently, we use the following constraint-based approach. First, we build a symbolic synthesis derivation, which may contain unknown refinement terms \(U_\Gamma\), and collect all subtyping, sharing, and transfer premises from the derivation into a system of typing constraints. Here \(\Delta\) records the desired sort of the unknown refinement term, and \(\Gamma\) records the context in which it must be well-formed. A solution to a system of typing constraints, is a map \(C : U \rightarrow \psi\) such that for every unknown \(U_\Gamma\), \(\Gamma \vdash \psi\) \(U \in \Delta\) and substituting \(\psi\) for \(U\) within the typing constraints yields valid subtyping, sharing, and transfer judgments.

Constraint solving To solve typing constraints, the algorithm first transforms them into validity constraints of one of two forms: \(\Gamma \models \psi \Rightarrow \psi'\) or \(\Gamma \models \psi \geq 0\); the interesting cases of this translation are shown in Fig. 9. Then, using the definition of validity (in the technical report [40]), we further reduce these into a system of:

1. Horn constraints of the form \(\psi_1 \land \ldots \land \psi_n \Rightarrow \psi_0\), and
2. resource constraints of the form \(\psi_1 \land \ldots \land \psi_n \Rightarrow \phi \geq 0\).

Here any \(\psi_i\) can be either a Boolean unknown \(U_\Gamma^B\) or a known refinement term, and \(\phi\) is a sum of zero or more numeric unknowns \(U_\Gamma^N\) and a known (linear) refinement term. While prior work has shown how to efficiently solve Horn constraints using predicate abstraction [51, 54], resource constraints present a new challenge, since they contain unknown terms of both Boolean and numeric sorts. In the interest of efficiency, our synthesis algorithm does not attempt to solve for both Boolean and numeric terms at the same time. Instead, it uses existing techniques to find a solution for the Horn constraints, and then plugs this solution into the resource constraints. Note that this approach does not sacrifice completeness, as long as the Horn solver returns the least-fixpoint (i.e. strongest) solution for each \(U_\Gamma^R\), since Boolean unknowns only appear negatively in resource constraints.

Resource constraints The main new challenge then is to solve a system of resource constraints of the form \(\psi \Rightarrow \phi \geq 0\), where \(\psi\) is now a known formula of the refinement logic. Since potential annotations in \(Re^2\) are restricted to linear terms over program variables, we can replace each unknown term \(U_\Gamma^R\) in \(\phi\) with a linear template \(\sum_{x \in X} C_i \cdot x\), where each \(C_i\) is an unknown integer coefficient and \(X\) is the set of all variables in \(\Gamma\) such that \(\Gamma \vdash x \in N\). After normalization, the system of resource constraints is reduced to the following doubly-quantified system of linear inequalities:

\[
\exists C_i, x, r : \bigwedge_{r \in K} r(X) \exists C_i, \sum_{x \in X} \bigwedge_{r \in K} r(X) \exists C_i, \sum_{x \in X}
\]

where each clause \(r\) is of the form \(\psi(X) \Rightarrow \sum_{x \in X} f(C_i \cdot x) \geq 0\), \(\psi\) is a known formula over the program variables \(X\), and each \(f\) is a linear function over unknown integer coefficients \(C_i\).

Note a crucial difference between these constraints and those generated by RaML: since RaML’s potential annotations are not dependent—i.e. \(r\) cannot mention program variables \(X\)—its resource constraints reduce to plain linear inequalities:

\[
\exists C_i, \sum_{x \in X} C_i \geq c (where c is a known constant), which can be handled by an LP solver. In our case, the challenge stems from the double quantification and the fact that individual clauses \(r\) are bounded by formulas \(\psi\), which are often nontrivial. For example, synthesizing the function range from Sec. 2 gives rise to the following (simplified) resource constraints:

\[
\exists C_0, \ldots, C_3, \forall a, b, v.
\]

Incremental solving Constraints of this form can be solved using counter-example guided synthesis (CEGIS) [62], which is, however, relatively expensive. We observe that in the context of synthesis we have to repeatedly solve similar systems of resource constraints because a program candidate is type-checked incrementally as it is being constructed, which corresponds to an incrementally growing set of clauses \(R\). Moreover,
we observe that as new clauses are added, only a few existing coefficients are typically invalidated, so we can avoid solving for all the coefficients from scratch. To this end, we develop an incremental version of the CEGIS algorithm, shown in Algorithm 1.

The goal of the algorithm is to find a solution \( C \cdot C_1 \rightarrow \mathcal{Z} \) that maps unknown coefficients to integers such that \( \forall x. R(\overline{C}, \overline{x}) \) holds (we write \( R(\overline{C}, \overline{x}) \) as a shorthand for \( \bigwedge_{r \in R} r(C, e) \)). The algorithm takes as input a set of clauses \( R \) (which includes both old and new clauses), the current solution \( C \) (new coefficients \( C_1 \) are mapped to 0) and the current set of examples \( E \), where an example \( e \in E \) is a partial assignment to universally-quantified variables \( e : X \rightarrow \mathbb{N} \).

The algorithm first queries the SMT solver for a counter-example \( e \) to the current solution. If no such counter-example exists, the solution is still valid (this happens surprisingly often, since many resource constraints are trivial). Otherwise, the current solution needs to be updated. To this end, a traditional CEGIS algorithm would query the SMT solver with the following synthesis constraint: \( \exists C_1. \bigwedge_{e \in E} R(C_1, e) \), which enforces that all clauses are satisfied on the extended set of examples. Instead, our incremental algorithm picks out only those clauses \( R' \) that are actually violated by the new counter-example; since in our setting \( R' \) is typically small, this optimization significantly reduces the size of the synthesis constraint and synthesis times for programs with dependent annotations (as we demonstrate in Sec. 5).

4.3 Implementation

We implemented the resource-guided synthesis algorithm in ReSyn, which extends SYNQUID with support for resource-annotated types and a resource constraint solver. Note that while our formalization is restricted to Booleans and length-indexed lists, our implementation supports the full expressiveness of SYNQUID’s types: types include integers and user-defined algebraic datatypes, and refinement formulas support sets and can mention arbitrary user-defined measures. More importantly, resource terms in ReSyn can mention integer variables and use subtraction, multiplication, conditional expressions, and numeric measures; finally, multiplicities on type variables can be dependent (mention variables). These changes have the following implications: (1) resource terms are not syntactically guaranteed to be non-negative, so we emit additional well-formedness constraints to enforce this; (2) resource terms are not syntactically restricted to be linear; our implementation is incomplete, and simply rejects the program if a nonlinear term arises; (3) subtyping and sharing constraints with conditional resource terms are decomposed into unconditional ones by moving the guard to the context, so the search space for all numeric unknowns remains unconditional; (4) to handle measure applications in resource constraints, we replace them with fresh integer variables, and avoid spurious counter-examples by explicitly instantiating the congruence axiom with all applications in the constraint.

5 Evaluation

We evaluated ReSyn using the following criteria:

- Relative performance: How do ReSyn’s synthesis times compare to SYNQUID’s? How much does the additional burden of solving resource constraints affect its performance?
- Efficacy of resource analysis: Can ReSyn discover more efficient programs than SYNQUID?
- Value of round-trip type checking: Does round-trip type checking afforded by the tight integration of resource analysis into SYNQUID effective at pruning the search space? How does it compare to the naive combination of synthesis and resource analysis?
- Value of incremental solving: To what extent does incremental solving of resource constraints improve ReSyn’s performance?

5.1 Relative Performance

To evaluate ReSyn’s performance relative to SYNQUID, we selected 43 problems from SYNQUID’s original suite, annotated them with resource bounds, and re-synthesized them with ReSyn. The rest of the original 64 benchmarks require non-linear bounds, and thus are out of scope of Re\(^2\). The details of this experiment are shown in Tab. 1, which compares ReSyn’s synthesis times against SYNQUID’s on these linear-bounded benchmarks.

Unsurprisingly, due to the additional constraint-solving, ReSyn generally performs worse than SYNQUID: the median synthesis time is about 2.5× higher. Note, however, that in return it provides provable guarantees about the performance of generated code. ReSyn was able to discover a more efficient implementation for only one of the original SYNQUID benchmarks (compress, discussed below). In general, these benchmarks contain only the minimal set of components required to produce a valid implementation, which makes it hard for SYNQUID to find a non-optimal version. Four of the
Table 1. Comparison of ReSyn and SynQiD. For each benchmark, we report the set of provided Components; cumulative size of synthesized Code (in AST nodes) for all goals; as well as running times (in seconds) for ReSyn (Time) and SynQiD (TimeNR).

<table>
<thead>
<tr>
<th>Group</th>
<th>Description</th>
<th>Components</th>
<th>Code</th>
<th>Time</th>
<th>TimeNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>List</td>
<td>is empty</td>
<td>true, false, =, #</td>
<td>16</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>duplicate each element</td>
<td>true, false, =, #</td>
<td>41</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>replicate</td>
<td>0, inc, dec, ≤, #</td>
<td>39</td>
<td>0.5</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>append two lists</td>
<td>0, inc, dec, ≤, #</td>
<td>31</td>
<td>2.9</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>take first n elements</td>
<td>0, inc, dec, ≤, #</td>
<td>38</td>
<td>1.5</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>drop first n elements</td>
<td>0, inc, dec, ≤, #</td>
<td>34</td>
<td>2.4</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>delete value</td>
<td>∈, #</td>
<td>49</td>
<td>3.3</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>zip</td>
<td>=, #</td>
<td>49</td>
<td>0.8</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>zip with</td>
<td>32</td>
<td>0.4</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>i-th element</td>
<td>0, inc, dec, ≤, #</td>
<td>35</td>
<td>0.5</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>index of element</td>
<td>0, inc, dec, ≤, #</td>
<td>30</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>insert at end</td>
<td>∈, #</td>
<td>43</td>
<td>0.5</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>balanced split</td>
<td>∈, #</td>
<td>42</td>
<td>0.4</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>reverse</td>
<td>∈, #</td>
<td>35</td>
<td>0.4</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>insert (sorted)</td>
<td>≤, #</td>
<td>57</td>
<td>2.0</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>extract minimum</td>
<td>≤, #</td>
<td>71</td>
<td>18.1</td>
<td>8.5</td>
</tr>
<tr>
<td></td>
<td>fold</td>
<td>≤, #</td>
<td>43</td>
<td>1.8</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>length using fold</td>
<td>0, inc, dec</td>
<td>39</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>append using fold</td>
<td>∈, #</td>
<td>42</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>map</td>
<td>∈, #</td>
<td>27</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td>Unique list</td>
<td>insert</td>
<td>∈, #</td>
<td>49</td>
<td>0.8</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>delete</td>
<td>∈, #</td>
<td>45</td>
<td>0.5</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>compressor</td>
<td>64</td>
<td>5.0</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>integer range</td>
<td>0, inc, dec, ≤, #</td>
<td>46</td>
<td>88.4</td>
<td>5.1</td>
</tr>
<tr>
<td></td>
<td>partition</td>
<td>≤, #</td>
<td>71</td>
<td>13.0</td>
<td>5.5</td>
</tr>
<tr>
<td>Sorted list</td>
<td>insert</td>
<td>∈, #</td>
<td>64</td>
<td>1.6</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>delete</td>
<td>∈, #</td>
<td>52</td>
<td>0.5</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>intersect</td>
<td>≤, #</td>
<td>71</td>
<td>17.0</td>
<td>0.8</td>
</tr>
<tr>
<td>Tree</td>
<td>node count</td>
<td>0, 1, ∈</td>
<td>34</td>
<td>3.8</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>preorder</td>
<td>∈, #</td>
<td>45</td>
<td>3.0</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>to list</td>
<td>∈, #</td>
<td>45</td>
<td>3.0</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>member</td>
<td>false, not, or, =</td>
<td>63</td>
<td>2.2</td>
<td>0.6</td>
</tr>
<tr>
<td>BST</td>
<td>member</td>
<td>true, false, ≤, #</td>
<td>72</td>
<td>0.5</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>insert</td>
<td>≤, #</td>
<td>90</td>
<td>4.5</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td>delete</td>
<td>≤, #</td>
<td>103</td>
<td>26.8</td>
<td>9.3</td>
</tr>
<tr>
<td></td>
<td>BST sort</td>
<td>≤, #</td>
<td>191</td>
<td>9.0</td>
<td>4.3</td>
</tr>
<tr>
<td>Binary Heap</td>
<td>insert</td>
<td>∈, #</td>
<td>90</td>
<td>3.2</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>member</td>
<td>false, not, or, ≤, #</td>
<td>78</td>
<td>2.3</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>1-element constructor</td>
<td>≤, #</td>
<td>44</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>2-element constructor</td>
<td>≤, #</td>
<td>91</td>
<td>0.7</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>3-element constructor</td>
<td>≤, #</td>
<td>274</td>
<td>21.4</td>
<td>4.0</td>
</tr>
</tbody>
</table>

benchmarks in Tab. 1 use advanced features of Re²: for example, any function using natural numbers to index or construct a data structure requires dependent potential annotations.

5.2 Case Studies

The value of resource-guided synthesis becomes clear when the library of components grows. To confirm this intuition, we assembled a suite of 16 case studies shown in Tab. 2, each exemplifying some feature of ReSyn.

Optimization The first six benchmarks showcase ReSyn’s ability to generate faster code than SynQiD (the cost metric in each case is the number of recursive calls). Benchmark 1 is triple from Sec. 2.3, where both SynQiD and ReSyn generate the same efficient solution; benchmark 2 is slight modification of this example: it uses a component append’, which traverses its second argument (unlike append, which traverses its first). In this case, ReSyn generates the efficient solution, associating the two calls to append’ to the left, while SynQiD still generates the same—now inefficient—solution, associating these calls to the right. In benchmark 3 ReSyn makes the optimal choice of accumulator to avoid a quadratic-time implementation. Benchmark 4 is compress from Tab. 1: the task is to remove adjacent duplicated from a list. Here SynQiD makes an unnecessary recursive call, resulting in a solution that is slightly shorter but runs in exponential time!

In other cases, ReSyn drastically changes the structure of the program to find an optimal implementation. Benchmark 5 is common from Sec. 2.1, where ReSyn must find an implementation that does not call member. Benchmark 6 works similarly, but computes the difference between two lists instead of their intersection. On these benchmarks, the performance disparity between ReSyn and SynQiD is much worse, as ReSyn must reject many more programs before it finds an appropriate implementation. On the other hand, these benchmarks also showcase the value of round-trip type checking: the column T-EAC reports synthesis times for a naive combination of synthesis and resource analysis, where we simply ask SynQiD to enumerate functionally correct programs until one type-checks under Re². As you can see, for benchmarks 5 and 6 this naive version times out after ten minutes.

Dependent Potentials Benchmarks 7–13 showcase fine-grained bounds that leverage dependent potential annotations. The first three of those synthesize a function insert that inserts an element into a sorted list. In benchmark 7 we use a simple linear bound (the length of the list), while benchmarks 8 and 9 specify a tighter bound: insert x xs can only make one recursive call per element of xs larger than x. These two examples showcase two different styles of specifying precise bounds: in 8 we define a custom measure n of list elements greater than a certain value; in 9, we instead annotate each list element with a conditional term indicating that it carries potential only if its value is larger than x. As discussed in Sec. 2, benchmark 13 (range) cannot be synthesized by SynQiD at all, because of restrictions on its termination checking mechanism, while ReSyn handles this benchmark out of the box.

For benchmarks 8–13, which make use of dependent potential annotations, we also report the synthesis times without incremental solving of resource constraints (T-Ninc), which are up to 2x higher.

Constant Resource As discussed in Sec. 3, a simple extension to Re² enables it to verify constant-resource implementations. We showcase this feature in benchmarks 14–16. Benchmark 15 is an example from [46], which compares a public list ys with a secret list zs. By allotting potential only to ys, we guarantee that the resource consumption of the generated
Table 2. Case Studies. For each synthesis problem, we report: the run time of ReSyn (T), SynQUD (T-NR), naive combination of SynQUD and resource analysis (T-EAC), ReSyn without incremental solving (T-NInc); as well as the tightest resource bound for the code generated by ReSyn (B) and by SynQUD (B-NR). Here, SL is the type of sorted lists, and CL refers to the type of lists without adjacent duplicates. TO is 10 min; all benchmarks count recursive calls.

### Table 2. Case Studies

<table>
<thead>
<tr>
<th>Description</th>
<th>Type Signature</th>
<th>Components</th>
<th>T</th>
<th>T-NR</th>
<th>T-EAC</th>
<th>T-NInc</th>
<th>B</th>
<th>B-NR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 triple</td>
<td>( \forall x. \text{List}(a) \to { \text{List}(a) } ) ( \forall x. \text{List}(a) \to { \text{List}(a) } )</td>
<td>append</td>
<td>0.9</td>
<td>0.4</td>
<td>0.4</td>
<td>-</td>
<td>(</td>
<td>x</td>
</tr>
<tr>
<td>2 triple</td>
<td>( \forall x. \text{List}(a) \to { \text{List}(a) } ) ( \forall x. \text{List}(a) \to { \text{List}(a) } )</td>
<td>append</td>
<td>2.8</td>
<td>0.4</td>
<td>1.2</td>
<td>-</td>
<td>(</td>
<td>x</td>
</tr>
<tr>
<td>3 concat list of lists</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>&amp;</td>
<td>3.2</td>
<td>0.9</td>
<td>1.1</td>
<td>-</td>
<td>(</td>
<td>x</td>
</tr>
<tr>
<td>4 compress</td>
<td>( \forall x. \text{List}(a) \to { \text{List}(a) } ) ( \forall x. \text{List}(a) \to { \text{List}(a) } )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 common</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 list difference</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 insert</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 insert'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 insert''</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 replace</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 take</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 drop</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 range</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14 CT insert</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15 CT compare</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16 compare</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Synthesis with Quantitative Objectives** Two lines of prior work on synthesis are explicitly concerned with optimizing resource usage. One is quantitative automata-theoretic synthesis, which has been used to synthesize optimal Mealy machines [8] and place synchronization in concurrent programs [11, 12, 28]. In contrast, we focus on synthesis of high-level programs that can manipulate custom data structures, which are out of reach for automata-theoretic synthesis.

The second relevant line of work is synthesis-aided compilation [49, 50, 56, 57]. This work is limited to generating low-level straight-line code, which is an easy target for correctness validation and cost estimation. Perhaps the closest work to ours is the Synapse tool [9], which supports a richer space of programs, but requires extensive guidance from the user (in the form of meta-sketches), and relies on bounded reasoning, which can only provide correctness and optimality guarantees for a finite set of inputs. In contrast, we use type-based verification and resource analysis techniques, which enable ReSyn to handle high-level recursive programs and provide guarantees for an unbounded set of inputs.

**Acknowledgments**

This article is based on research supported by the United States Air Force under DARPA AA Contract FA8750-18-C-0092 and DARPA STAC Contract FA8750-15-C-0082, and by the National Science Foundation under SaTC Award 1801369, SHF Award 1812876, and CAREER Award 1845514. Any opinions, findings, and conclusions contained in this document are those of the authors and do not necessarily reflect the views of the sponsoring organizations.

**References**

Resource-Guided Program Synthesis

18th Conference (LPAR ’12), 1–11.


Langs. (POPL '10), 223–236.


