Relational Analysis of (Co)inductive Predicates, (Co)algebraic Datatypes, and (Co)recursive Functions[⋆]

Jasmin Christian Blanchette

Institut für Informatik, Technische Universität München, Germany
`blanchette@in.tum.de`

Abstract. This paper presents techniques for applying a finite relational model finder to logical specifications that involve (co)inductive predicates, (co)algebraic datatypes, and (co)recursive functions. In contrast to previous work, which focused on algebraic datatypes and restricted occurrences of unbounded quantifiers in formulas, we can handle arbitrary formulas by means of a three-valued Kleene logic. The techniques form the basis of the counterexample generator Nitpick for Isabelle/HOL. As a case study, we consider a coalgebraic lazy list type.

1 Introduction

SAT and SMT solvers, model checkers, model finders, and other lightweight formal methods are today available to test or verify specifications written in various languages. These tools are often integrated in more powerful systems, such as interactive theorem provers, to discharge proof obligations or generate (counter)models.

For testing logical specifications, a particularly attractive approach is to express these in first-order relational logic (FORL) and use a model finder such as Kodkod [30] to find counterexamples. FORL extends traditional first-order logic (FOL) with relational calculus operators and the transitive closure, and offers a good compromise between automation and expressiveness. Kodkod relies on a SAT solver and forms the basis of Alloy [15]. In a case study, the Alloy Analyzer checked a mechanized version of the paper proof of the Mondex protocol and revealed several bugs in the proof [27].

However, FORL lacks the high-level definitional principles usually provided in interactive theorem provers, namely (co)inductive predicates, (co)algebraic datatypes, and (co)recursive functions (Sect. 3). Solutions have been proposed by Kuncak and Jackson [21], who modeled lists and trees in Alloy, and Dunets et al. [10], who showed how to translate algebraic datatypes and recursive functions in the context of the first-order theorem prover KIV. In both cases, the translation is restricted to formulas whose prenex normal forms contain no unbounded universal quantifiers ranging over datatypes.

This paper generalizes previous work in several directions: First, we lift the unbounded quantifier restriction by using a three-valued logic coded in terms of the binary logic FORL (Sect. 4.2). Second, we show how to translate (co)inductive predicates, coalgebraic datatypes, and corecursive functions (Sect. 5). Third, in our treatment of algebraic datatypes, we show how to handle mutually recursive datatypes (Sect. 5.2).

The use of a three-valued Kleene logic makes it possible to analyze formulas such as $\mathit{True} \vee \forall n^{\mathit{nat}}.\ P(n)$, which are rejected by Kuncak and Jackson's syntactic criterion.

[⋆] This work is supported by the DFG grant Ni 491/11-1.

Unbounded universal quantification is still problematic in general, but suitable definitional principles and their proper handling mitigate this problem.

The ideas presented here form the basis of the higher-order counterexample generator Nitpick [5], which is included with recent versions of Isabelle/HOL [25]. As a case study, we employ Nitpick on a small theory of coalgebraic (lazy) lists (Sect. 6). To simplify the presentation, we use FOL as our specification language in the paper; issues specific to higher-order logic (HOL) are mostly orthogonal and covered elsewhere [5].

2 Logics

2.1 First-Order Logic (FOL)

The first-order logic that will serve as our specification language is essentially the first-order fragment of HOL [7, 12]. The types and terms are given below.

Types:		*Terms:*	
$\sigma ::= \kappa$	(atomic type)	$t ::= x^{\sigma}$	(variable)
$\quad\mid \alpha$	(type variable)	$\quad\mid c^{\tau}(t,\ldots,t)$	(function term)
$\tau ::= (\sigma,\ldots,\sigma) \to \sigma$	(function type)	$\quad\mid \forall x^{\sigma}.\, t$	(universal quantification)

The standard semantics interprets the Boolean type o and the constants False^{o}, True^{o}, $\longrightarrow^{(o,o)\to o}$ (implication), $\simeq^{(\sigma,\sigma)\to o}$ (equality on basic type σ), and $\mathit{if\ then\ else}^{(o,\sigma,\sigma)\to\sigma}$. Formulas are terms of type o. We assume throughout this paper that terms are well-typed using the standard typing rules and usually omit the type superscripts. In conformity with first-order practice, application of x and y on f is written $f(x,y)$, the function type $() \to \sigma$ is identified with σ, and the parentheses in the function term $c()$ are optional. We also assume that the connectives $\neg$, $\wedge$, $\vee$ and existential quantification are available.

In contrast to HOL, our logic requires variables to range over basic types, and it forbids partial function application and λ-abstractions. On the other hand, it supports the limited form of polymorphism provided by proof assistants for HOL [14, 25, 29], with the restriction that type variables may only be instantiated by atomic types (or left uninstantiated in a polymorphic formula).

Types and terms are interpreted in the standard set-theoretic way, relative to a scope that fixes the interpretation of basic types. A *scope* S is a function from basic types to nonempty sets (domains), which need not be finite.[1] We require $S(o) = \{\mathit{ff}, \mathit{tt}\}$.

The standard interpretation $[\![\tau]\!]_S$ of a type τ is given by

$$[\![\sigma]\!]_S = S(\sigma) \qquad [\![(\sigma_1,\ldots,\sigma_n) \to \sigma]\!]_S = [\![\sigma_1]\!]_S \times \cdots \times [\![\sigma_n]\!]_S \to [\![\sigma]\!]_S,$$

where $A \to B$ denotes the set of (total) functions from A to B. In contexts where S is clear, the cardinality of $[\![\tau]\!]_S$ is written $|\tau|$.

2.2 First-Order Relational Logic (FORL)

Our analysis logic, first-order relational logic, combines elements from FOL and relational calculus extended with the transitive closure [15, 30]. Formulas involve variables and terms ranging over relations (sets of tuples drawn from a universe of atoms) of arbitrary arities. The logic is unsorted, but each term denotes a relation of a fixed arity.

[1] The use of the word "scope" for a domain specification is consistent with Jackson [15].

Formulas:		*Terms:*	
φ ::= false	(falsity)	r ::= none	(empty set)
$\mid \mathsf{true}$	(truth)	$\mid \mathsf{iden}$	(identity relation)
$\mid m\ r$	(multiplicity constraint)	$\mid \mathsf{a}_i$	(atom)
$\mid r \simeq r$	(equality)	$\mid x$	(variable)
$\mid r \subseteq r$	(inclusion)	$\mid r^+$	(transitive closure)
$\mid \neg\,\varphi$	(negation)	$\mid r.r$	(dot-join)
$\mid \varphi \wedge \varphi$	(conjunction)	$\mid r \times r$	(Cartesian product)
$\mid \forall x{\in}r\colon \varphi$	(universal quantification)	$\mid r \cup r$	(union)
		$\mid r - r$	(difference)
m ::= $\mathsf{no} \mid \mathsf{lone} \mid \mathsf{one} \mid \mathsf{some}$		$\mid \mathsf{if}\ \varphi\ \mathsf{then}\ r\ \mathsf{else}\ r$	(conditional)

The universe of discourse is $\mathscr{A} = \{\mathsf{a}_1, \ldots, \mathsf{a}_k\}$, where each a_i is an uninterpreted atom. Atoms and n-tuples are identified with singleton sets and singleton n-ary relations, respectively. Bound variables in quantifications range over the tuples in a relation; thus, $\forall x \in (\mathsf{a}_1 \cup \mathsf{a}_2) \times \mathsf{a}_3\colon \varphi(x)$ is equivalent to $\varphi(\mathsf{a}_1 \times \mathsf{a}_3) \wedge \varphi(\mathsf{a}_2 \times \mathsf{a}_3)$.

Although they are not listed above, we will sometimes make use of $\vee$, $\longrightarrow$, *, and $\cap$ as well. The constraint $\mathsf{no}\ r$ expresses that r is the empty relation, $\mathsf{one}\ r$ expresses that r is a singleton, $\mathsf{lone}\ r \iff \mathsf{no}\ r \vee \mathsf{one}\ r$, and $\mathsf{some}\ r \iff \neg\,\mathsf{no}\ r$. The dot-join operator is unconventional; its semantics is given by the equation

$$[\![r.s]\!] = \{(r_1, \ldots, r_{m-1}, s_2, \ldots, s_n) \mid \exists t.\ (r_1, \ldots, r_{m-1}, t) \in [\![r]\!] \wedge (t, s_2, \ldots, s_n) \in [\![s]\!]\}.$$

The operator admits three important special cases. Let s be unary and r, r' be binary relations. The expression $s.r$ gives the direct image of the set s under r; if s is a singleton and r a function, it coincides with the function application $r(s)$. Analogously, $r.s$ gives the inverse image of s under r. Finally, $r.r'$ expresses relational composition.

The following FORL specification attempts to fit 30 pigeons in 29 holes:

$$\begin{array}{l}
\textbf{vars}\ \mathit{pigeons} = \{\mathsf{a}_1, \ldots, \mathsf{a}_{30}\},\ \mathit{holes} = \{\mathsf{a}_{31}, \ldots, \mathsf{a}_{59}\} \\
\textbf{var}\ \emptyset \subseteq \mathit{nest} \subseteq \{\mathsf{a}_1, \ldots, \mathsf{a}_{30}\} \times \{\mathsf{a}_{31}, \ldots, \mathsf{a}_{59}\} \\
\textbf{solve}\ (\forall p \in \mathit{pigeons}\colon \mathsf{one}\ p.\mathit{nest}) \wedge (\forall h \in \mathit{holes}\colon \mathsf{lone}\ \mathit{nest}.h)
\end{array}$$

The variables *pigeons* and *holes* are given fixed values, whereas *nest* is specified with a lower and an upper bound. The constraint $\mathsf{one}\ p.\mathit{nest}$ states that pigeon p is in relation with exactly one hole, and $\mathsf{lone}\ \mathit{nest}.h$ that hole h is in relation with at most one pigeon. Taken as a whole, the formula states that *nest* is a one-to-one function. It is, of course, not satisfiable, a fact that Kodkod can establish in less than a second.

3 Definitional Principles

3.1 Simple Definitions

We extend our specification logic FOL with several definitional principles to introduce new constants and types. The first principle defines a constant as equal to another term:

$$\textbf{definition}\ c^{\tau}\ \textbf{where}\ c(\bar{x}) \simeq t$$

Logically, the above definition is equivalent to the axiom $\forall \bar{x}.\ c(\bar{x}) \simeq t$.

Provisos: The constant c is fresh, the variables $\bar{x}$ are distinct, and the right-hand side t does not refer to any other free variables than $\bar{x}$, to any undefined constants or c, or to any type variables not occurring in τ. These restrictions ensure consistency [32].

3.2 (Co)inductive Predicates

The **inductive** and **coinductive** commands define inductive and coinductive predicates specified by their introduction rules:

$$\begin{array}{l}
[\textbf{co}]\textbf{inductive}\ p^{\tau}\ \textbf{where} \\
p(\bar{t}_{11}) \wedge \cdots \wedge p(\bar{t}_{1\ell_1}) \wedge Q_1 \longrightarrow p(\bar{u}_1) \\
\quad\vdots \\
p(\bar{t}_{n1}) \wedge \cdots \wedge p(\bar{t}_{n\ell_n}) \wedge Q_n \longrightarrow p(\bar{u}_n)
\end{array}$$

Provisos: The constant p is fresh, and the arguments to p and the side conditions Q_i do not refer to p, undeclared constants, or any type variables not occurring in τ.

The introduction rules may involve any number of free variables $\bar{y}$. The syntactic restrictions on the rules ensure monotonicity; by the Knaster–Tarski theorem, the fixed point equation

$$p(\bar{x}) \ \simeq\ \exists\bar{y}.\ \bigvee_{j=1}^{n} \bar{x} \simeq \bar{u}_j \wedge p(\bar{t}_{j1}) \wedge \cdots \wedge p(\bar{t}_{j\ell_j}) \wedge Q_j$$

admits a least and a greatest solution [13, 26]. Inductive definitions provide the least fixed point, and coinductive definitions provide the greatest fixed point.

As an example, assuming a type *nat* of natural numbers generated freely by 0^{nat} and $Suc^{nat\to nat}$, the following definition introduces the predicate *even* of even numbers:

$$\begin{array}{l}
\textbf{inductive}\ even^{nat\to o}\ \textbf{where} \\
even(0) \\
even(n) \longrightarrow even(Suc(Suc(n)))
\end{array}$$

The associated fixed point equation is

$$even(x) \ \simeq\ \exists n.\ x \simeq 0 \vee x \simeq Suc(Suc(n)) \wedge even(n).$$

The syntax can be generalized to support mutual definitions, as in the next example:

$$\begin{array}{l}
\textbf{inductive}\ even^{nat\to o}\ \textbf{and}\ odd^{nat\to o}\ \textbf{where} \\
even(0) \\
even(n) \longrightarrow odd(Suc(n)) \\
odd(n) \longrightarrow even(Suc(n))
\end{array}$$

Mutual definitions for $p_1, \ldots, p_m$ can be reduced to a single predicate q whose domain is the disjoint sum of the domains of each p_i [26]. Assuming *Inl* and *Inr* are the disjoint sum constructors, the definition of *even* and *odd* is replaced by

$$\begin{array}{l}
\textbf{inductive}\ even\_or\_odd^{(nat,nat)\,sum\to o}\ \textbf{where} \\
even\_or\_odd(Inl(0)) \\
even\_or\_odd(Inl(n)) \longrightarrow even\_or\_odd(Inr(Suc(n))) \\
even\_or\_odd(Inr(n)) \longrightarrow even\_or\_odd(Inl(Suc(n)))
\end{array}$$

$$\begin{array}{l}
\textbf{definition}\ even^{nat\to o}\ \textbf{where}\ even(n) \simeq even\_or\_odd(Inl(n)) \\
\textbf{definition}\ odd^{nat\to o}\ \textbf{where}\ odd(n) \simeq even\_or\_odd(Inr(n))
\end{array}$$

3.3 (Co)algebraic Datatypes

The **datatype** and **codatatype** commands define mutually recursive (co)algebraic datatypes specified by their constructors:

$$\begin{array}{l} [\textbf{co}]\textbf{datatype}\ (\bar{\alpha})\,\kappa_1 = C_{11}\ \big[\textbf{of}\ \bar{\sigma}_{11}\big]\ \mid\ \cdots\ \mid\ C_{1\ell_1}\ \big[\textbf{of}\ \bar{\sigma}_{1\ell_1}\big] \\ \qquad\quad \textbf{and}\ \ldots \\ \qquad\quad \textbf{and}\ (\bar{\alpha})\,\kappa_n = C_{n1}\ \big[\textbf{of}\ \bar{\sigma}_{n1}\big]\ \mid\ \cdots\ \mid\ C_{n\ell_n}\ \big[\textbf{of}\ \bar{\sigma}_{n\ell_n}\big] \end{array}$$

The defined types $(\bar{\alpha})\,\kappa_i$ are parameterized by a list of distinct type variables $\bar{\alpha}$, providing type polymorphism. Each constructor C_{ij} has type $\bar{\sigma}_{ij} \to (\bar{\alpha})\,\kappa_i$.

Provisos: The type names κ_i and the constructor constants C_{ij} are fresh and distinct, the type parameters $\bar{\alpha}$ are distinct, and the argument types $\bar{\sigma}_{ij}$ do not refer to any other type variables than $\bar{\alpha}$ (but may refer to the types $(\bar{\alpha})\,\kappa_i$ being defined).

The commands can be used to define natural numbers, pairs, finite lists, and possibly infinite lazy lists as follows:

$$\begin{array}{ll} \textbf{datatype}\ \mathit{nat} = 0 \mid \mathit{Suc}\ \textbf{of}\ \mathit{nat} & \textbf{datatype}\ \alpha\,\mathit{list} = \mathit{Nil} \mid \mathit{Cons}\ \textbf{of}\ (\alpha, \alpha\,\mathit{list}) \\ \textbf{datatype}\ (\alpha,\beta)\,\mathit{pair} = \mathit{Pair}\ \textbf{of}\ (\alpha,\beta) & \textbf{codatatype}\ \alpha\,\mathit{llist} = \mathit{LNil} \mid \mathit{LCons}\ \textbf{of}\ (\alpha, \alpha\,\mathit{llist}) \end{array}$$

Defining a (co)datatype introduces the appropriate axioms for the constructors [26]. It also introduces the syntax $\mathit{case}\ t\ \mathit{of}\ C_{i1}(\bar{x}_1) \Rightarrow u_1 \mid \ldots \mid C_{i\ell_i}(\bar{x}_{\ell_i}) \Rightarrow u_{\ell_i}$, characterized by $\forall \bar{x}_j.\ (\mathit{case}\ C_{ij}(\bar{x}_j)\ \mathit{of}\ C_{i1}(\bar{x}_1) \Rightarrow u_1 \mid \ldots \mid C_{i\ell_i}(\bar{x}_{\ell_i}) \Rightarrow u_{\ell_i}) \simeq u_j$ for $j \in \{1, \ldots, \ell_i\}$.

3.4 (Co)recursive Functions

The **primrec** command defines primitive recursive functions on algebraic datatypes:

$$\begin{array}{lcl} \multicolumn{3}{l}{\textbf{primrec}\ f_1^{\tau_1}\ \textbf{and}\ \ldots\ \textbf{and}\ f_n^{\tau_n}\ \textbf{where}} \\ f_1(C_{11}(\bar{x}_{11}), \bar{z}_{11}) \simeq t_{11} & \ldots & f_1(C_{1\ell_1}(\bar{x}_{1\ell_1}), \bar{z}_{1\ell_1}) \simeq t_{1\ell_1} \\ \quad\vdots & & \\ f_n(C_{n1}(\bar{x}_{n1}), \bar{z}_{n1}) \simeq t_{n1} & \ldots & f_n(C_{n\ell_n}(\bar{x}_{n\ell_n}), \bar{z}_{n\ell_n}) \simeq t_{n\ell_n} \end{array}$$

Provisos: The constants f_i are fresh and distinct, the variables $\bar{x}_{ij}$ and $\bar{z}_{ij}$ are distinct for any given i and j, the right-hand sides t_{ij} involve no other variables than $\bar{x}_{ij}$ and $\bar{z}_{ij}$ and no type variables that do not occur in τ_i, and the first argument of any recursive call must be one of the $\bar{x}_{ij}$'s. The recursion is well-founded because each recursive call peels off one constructor from the first argument.

Corecursive function definitions follow a rather different syntactic schema, with a single equation per function f_i that must return type $(\bar{\alpha})\,\kappa_i$:

$$\begin{array}{l} \textbf{coprimrec}\ f_1^{\tau_1}\ \textbf{and}\ \ldots\ \textbf{and}\ f_n^{\tau_n}\ \textbf{where} \\ f_1(\bar{y}_1)\ \simeq\ \mathit{if}\ Q_{11}\ \mathit{then}\ t_{11}\ \mathit{else}\ \mathit{if}\ Q_{12}\ \mathit{then}\ \ldots\ \mathit{else}\ t_{1\ell_1} \\ \quad\vdots \\ f_n(\bar{y}_n)\ \simeq\ \mathit{if}\ Q_{n1}\ \mathit{then}\ t_{n1}\ \mathit{else}\ \mathit{if}\ Q_{n2}\ \mathit{then}\ \ldots\ \mathit{else}\ t_{n\ell_n} \end{array}$$

Provisos: The constants f_i are fresh and distinct, the variables $\bar{y}_i$ are distinct, the right-hand sides involve no other variables than $\bar{y}_i$, no corecursive calls occur in the conditions

Q_{ij}, and either t_{ij} does not involve any corecursive calls or it has the form $C_{ij}(\bar{u}_{ij})$.[2] The syntax can be relaxed to allow a *case* expression instead of a sequence of conditionals.

The following examples define concatenation for α *list* and α *llist*:

$$\textsf{\textbf{primrec}}\ cat^{(\alpha\,list,\,\alpha\,list)\to\alpha\,list}\ \textsf{\textbf{where}}$$
$$cat(Nil, zs) \simeq zs \qquad cat(Cons(y, ys), zs) \simeq Cons(y, cat(ys, zs))$$

$$\textsf{\textbf{coprimrec}}\ lcat^{(\alpha\,llist,\,\alpha\,llist)\to\alpha\,llist}\ \textsf{\textbf{where}}$$
$$lcat(ys, zs) \simeq case\ ys\ of\ LNil \Rightarrow zs \mid LCons(y, ys') \Rightarrow LCons(y, lcat(ys', zs))$$

4 Basic Translations

4.1 A Sound and Complete Translation

This section presents the translation of FOL to FORL, excluding the definitional principles from Sect. 3. We consider only finite domains; for these the translation is sound and complete. We start by mapping FOL types τ to sets of FORL atom tuples $\langle\!\langle\tau\rangle\!\rangle$:

$$\langle\!\langle\sigma\rangle\!\rangle = \{\mathsf{a}_1, \ldots, \mathsf{a}_{|\sigma|}\} \qquad \langle\!\langle(\sigma_1, \ldots, \sigma_n) \to \sigma\rangle\!\rangle = \langle\!\langle\sigma_1\rangle\!\rangle \times \cdots \times \langle\!\langle\sigma_n\rangle\!\rangle \times \langle\!\langle\sigma\rangle\!\rangle.$$

For simplicity, we reuse the same atoms for distinct basic types. A real implementation can benefit from using distinct atoms because it facilitates symmetry breaking [30].

For each free variable or nonstandard constant u^τ, we generate the bounds declaration **var** $\emptyset \subseteq u \subseteq \langle\!\langle\tau\rangle\!\rangle$ as well as a constraint $\Phi(u)$ to ensure that single values are singletons and functions are functions:

$$\Phi(u^\sigma) = \textsf{one}\ u \qquad \Phi(u^{(\varsigma_1,\ldots,\varsigma_n)\to\varsigma}) = \forall x_1 \in \langle\!\langle\varsigma_1\rangle\!\rangle, \ldots, x_n \in \langle\!\langle\varsigma_n\rangle\!\rangle\colon \textsf{one}\ x_n.(\ldots.(x_1.u)\ldots).$$

Since FORL distinguishes between formulas and terms, the translation to FORL is performed by two mutually recursive functions, $\mathsf{F}\langle\!\langle t\rangle\!\rangle$ and $\mathsf{T}\langle\!\langle t\rangle\!\rangle$:[3]

$$\begin{array}{rlrl}
\mathsf{F}\langle\!\langle False\rangle\!\rangle &= \textsf{false} & \mathsf{T}\langle\!\langle x\rangle\!\rangle &= x \\
\mathsf{F}\langle\!\langle True\rangle\!\rangle &= \textsf{true} & \mathsf{T}\langle\!\langle False\rangle\!\rangle &= \mathsf{a}_1 \\
\mathsf{F}\langle\!\langle t \simeq u\rangle\!\rangle &= \mathsf{T}\langle\!\langle t\rangle\!\rangle \simeq \mathsf{T}\langle\!\langle u\rangle\!\rangle & \mathsf{T}\langle\!\langle True\rangle\!\rangle &= \mathsf{a}_2 \\
\mathsf{F}\langle\!\langle t \longrightarrow u\rangle\!\rangle &= \mathsf{F}\langle\!\langle t\rangle\!\rangle \longrightarrow \mathsf{F}\langle\!\langle u\rangle\!\rangle & \mathsf{T}\langle\!\langle if\ t\ then\ u_1\ else\ u_2\rangle\!\rangle &= \textsf{if}\ \mathsf{F}\langle\!\langle t\rangle\!\rangle\ \textsf{then}\ \mathsf{T}\langle\!\langle u_1\rangle\!\rangle\ \textsf{else}\ \mathsf{T}\langle\!\langle u_2\rangle\!\rangle \\
\mathsf{F}\langle\!\langle \forall x^\sigma\!.\, t\rangle\!\rangle &= \forall x \in \langle\!\langle\sigma\rangle\!\rangle\colon \mathsf{F}\langle\!\langle t\rangle\!\rangle & \mathsf{T}\langle\!\langle c(t_1, \ldots, t_n)\rangle\!\rangle &= \mathsf{T}\langle\!\langle t_n\rangle\!\rangle.(\ldots.(\mathsf{T}\langle\!\langle t_1\rangle\!\rangle.c)\ldots) \\
\mathsf{F}\langle\!\langle t\rangle\!\rangle &= \mathsf{T}\langle\!\langle t\rangle\!\rangle \simeq \mathsf{T}\langle\!\langle True\rangle\!\rangle & \mathsf{T}\langle\!\langle t^o\rangle\!\rangle &= \mathsf{T}\langle\!\langle if\ t\ then\ True\ else\ False\rangle\!\rangle.
\end{array}$$

The metavariable c ranges over nonstandard constants, so that the $\mathsf{T}\langle\!\langle t^o\rangle\!\rangle$ equation is used for $\simeq$ and $\longrightarrow$ (as well as for $\forall$). The Boolean values false and true are arbitrarily coded as a_1 and a_2 when they appear as FORL terms.

Theorem 4.1. *The FOL formula P with free variables and nonstandard constants $u_1^{\tau_1}$, $\ldots$, $u_n^{\tau_n}$ is satisfiable for a given finite scope iff the FORL formula $\mathsf{F}\langle\!\langle P\rangle\!\rangle \wedge \bigwedge_{j=1}^n \Phi(u_j)$ with bounds $\emptyset \subseteq u_j \subseteq \langle\!\langle\tau_j\rangle\!\rangle$ is satisfiable for the same scope.*

[2] Other authors formulate corecursion in terms of selectors instead of constructors [16].

[3] Metatheoretic functions here and elsewhere are defined using sequential pattern matching.

Proof. Let $[\![t]\!]_M$ denote the set-theoretic semantics of the FOL term t w.r.t. a model M and the given scope S, let $[\![\varphi]\!]_V$ denote the truth value of the FORL formula φ w.r.t. a variable valuation V and the scope S, and let $[\![r]\!]_V$ denote the set-theoretic semantics of the FORL term r w.r.t. V and S. Furthermore, for $v \in [\![\sigma]\!]_S$, let $\lfloor v \rfloor$ denote the corresponding value in $\langle\!\langle \sigma \rangle\!\rangle$, with $\lfloor ff \rfloor = \mathsf{a}_1$ and $\lfloor tt \rfloor = \mathsf{a}_2$. Using recursion induction, it is straightforward to prove that $[\![\mathsf{F}\langle\!\langle t^o \rangle\!\rangle]\!]_V \iff [\![t]\!]_M = tt$ and $[\![\mathsf{T}\langle\!\langle t \rangle\!\rangle]\!]_V = \lfloor [\![t]\!]_M \rfloor$ if $V(u_j) = \lfloor M(u_j) \rfloor$ for all u_j's. Moreover, from a satisfying valuation V of the u_j's, we can construct a FOL model M such that $\lfloor M(u_j) \rfloor = V(u_j)$; the Φ constraints and the bounds ensure that such a model exists. Hence, $[\![\mathsf{F}\langle\!\langle P \rangle\!\rangle]\!]_V \iff [\![P]\!]_M = tt$. □

The translation is parameterized by a scope, which specifies the exact cardinalities of the basic types occurring in the formula. To exhaust all models up to a cardinality bound k for n basic types, a model finder must a priori iterate through k^n combinations of cardinalities and must consider all models for each of these combinations. This can be made more efficient by taking the cardinalities as upper bounds rather than exact bounds (Alloy's default mode of operation [15, p. 129]) or by inferring scope monotonicity [4,21].

4.2 Approximation of Infinite Types and Partiality

Besides its lack of support for the definitional principles, the above translation suffers from a serious limitation: It disregards infinite types such as natural numbers, lists, and trees, which are ubiquitous in real-world specifications. Fortunately, it is not hard to adapt the translation to take these into account in a sound (but incomplete) way.

Given an infinite atomic type κ, we consider a finite subset of $[\![\kappa]\!]_S$ and map every element not in this subset to a special undefined value $\bot$. For the type *nat* of natural numbers, an obvious choice is to consider prefixes $\{0, \ldots, K\}$ of $\mathbb{N}$ and map numbers $> K$ to $\bot$. Observe that the successor function *Suc* becomes partial, with $Suc\ K = \bot$. The technique can also be used to speed up the analysis of finite types with a high cardinality: We can approximate a 256-value *byte* type by a subset of, say, 5 values.

Leaving out some elements of atomic types means that we must cope with partiality. Not only may functions be partial, but any term or formula can evaluate to $\bot$. The logic becomes a three-valued Kleene logic [17]. Universal quantifiers whose bound variable ranges over an approximated type, such as $\forall n^{nat}.\ P(n)$, will evaluate to either *False* (if $P(n)$ gives *False* for some $n \le K$) or $\bot$, but never to *True*, since we do not know whether $P(K+1)$, $P(K+2)$, ..., are true.

Partiality can be encoded in FORL as follows. Inside terms, we let none (the empty set) stand for $\bot$. This choice is convenient because none is an absorbing element for the dot-join operator, which models function application; thus, $f(\bot) = \bot$. Inside a formula, we keep track of the polarity of the subformulas: In positive contexts (i.e., under an even number of negations), true codes *True* and false codes *False* or $\bot$; in negative contexts, false codes *False* and true codes *True* or $\bot$.

The translation of FOL terms is performed by two functions, $\mathsf{F}^s\langle\!\langle t \rangle\!\rangle$ and $\mathsf{T}\langle\!\langle t \rangle\!\rangle$, where s indicates the polarity ($+$ or $-$):

$$\mathsf{F}^s\langle\!\langle False \rangle\!\rangle = \mathsf{false} \qquad \mathsf{T}\langle\!\langle x \rangle\!\rangle = x$$
$$\mathsf{F}^s\langle\!\langle True \rangle\!\rangle = \mathsf{true} \qquad \mathsf{T}\langle\!\langle False \rangle\!\rangle = \mathsf{a}_1$$

$$
\begin{aligned}
\mathsf{F}^{+}\langle\!\langle t \simeq u\rangle\!\rangle &= \mathsf{some}\ (\mathsf{T}\langle\!\langle t\rangle\!\rangle \cap \mathsf{T}\langle\!\langle u\rangle\!\rangle) \\
\mathsf{F}^{-}\langle\!\langle t \simeq u\rangle\!\rangle &= \mathsf{lone}\ (\mathsf{T}\langle\!\langle t\rangle\!\rangle \cup \mathsf{T}\langle\!\langle u\rangle\!\rangle) \\
\mathsf{F}^{s}\langle\!\langle t \longrightarrow u\rangle\!\rangle &= \mathsf{F}^{-s}\langle\!\langle t\rangle\!\rangle \longrightarrow \mathsf{F}^{s}\langle\!\langle u\rangle\!\rangle \\
\mathsf{F}^{+}\langle\!\langle \forall x^{\sigma}.\, t\rangle\!\rangle &= \mathsf{false} \quad \text{if } |\langle\!\langle \sigma\rangle\!\rangle| < |\sigma| \\
\mathsf{F}^{s}\langle\!\langle \forall x^{\sigma}.\, t\rangle\!\rangle &= \forall x \in \langle\!\langle \sigma\rangle\!\rangle\text{: } \mathsf{F}^{s}\langle\!\langle t\rangle\!\rangle \\
\mathsf{F}^{+}\langle\!\langle t\rangle\!\rangle &= \mathsf{T}\langle\!\langle t\rangle\!\rangle \simeq \mathsf{T}\langle\!\langle \mathit{True}\rangle\!\rangle \\
\mathsf{F}^{-}\langle\!\langle t\rangle\!\rangle &= \mathsf{T}\langle\!\langle t\rangle\!\rangle \not\simeq \mathsf{T}\langle\!\langle \mathit{False}\rangle\!\rangle
\end{aligned}
\qquad
\begin{aligned}
\mathsf{T}\langle\!\langle \mathit{True}\rangle\!\rangle &= \mathsf{a}_2 \\
\mathsf{T}\langle\!\langle \mathit{if}\ t\ \mathit{then}\ u_1\ \mathit{else}\ u_2\rangle\!\rangle &= \mathsf{if}\ \mathsf{F}^{+}\langle\!\langle t\rangle\!\rangle\ \mathsf{then}\ \mathsf{T}\langle\!\langle u_1\rangle\!\rangle \\
&\quad\ \mathsf{else\ if}\ \neg\,\mathsf{F}^{-}\langle\!\langle t\rangle\!\rangle\ \mathsf{then}\ \mathsf{T}\langle\!\langle u_2\rangle\!\rangle \\
&\quad\ \mathsf{else\ none} \\
\mathsf{T}\langle\!\langle c(t_1,\ldots,t_n)\rangle\!\rangle &= \mathsf{T}\langle\!\langle t_n\rangle\!\rangle.(\ldots.(\mathsf{T}\langle\!\langle t_1\rangle\!\rangle.c)\ldots) \\
\mathsf{T}\langle\!\langle t^{o}\rangle\!\rangle &= \mathsf{T}\langle\!\langle \mathit{if}\ t\ \mathit{then}\ \mathit{True} \\
&\qquad\quad \mathit{else}\ \mathit{False}\rangle\!\rangle.
\end{aligned}
$$

In the equation for implication, $-s$ denotes $-$ if s is $+$ and $+$ if s is $-$. Taken together, $(\mathsf{F}^{+}\langle\!\langle t\rangle\!\rangle, \mathsf{F}^{-}\langle\!\langle t\rangle\!\rangle)$ encode a three-valued logic, with $(\mathsf{true}, \mathsf{true})$ corresponding to *True*, $(\mathsf{false}, \mathsf{true})$ corresponding to $\bot$, and $(\mathsf{false}, \mathsf{false})$ corresponding to *False*. The case $(\mathsf{true}, \mathsf{false})$ is impossible by construction.

When mapping FOL types to sets of FORL atom tuples, basic types σ are now allowed to take any finite cardinality $|\langle\!\langle \sigma\rangle\!\rangle| \le |\sigma|$. We also need to relax the definition of $\Phi(u)$ to allow empty sets, by substituting lone for one.

Theorem 4.2. *Given a FOL formula P with free variables and nonstandard constants* $u_1^{\tau_1}, \ldots, u_n^{\tau_n}$ *and a scope S, the FORL formula* $\mathsf{F}^{+}\langle\!\langle P\rangle\!\rangle \wedge \bigwedge_{j=1}^{n} \Phi(u_j)$ *with bounds* $\emptyset \subseteq u_j \subseteq \langle\!\langle \tau_j\rangle\!\rangle$ *is satisfiable for S only if P is satisfiable for S.*

Proof. The proof is similar to that of Theorem 4.1, but partiality requires us to compare the actual value of a FORL expression with its expected value using $\subseteq$ rather than $=$. Using recursion induction, we can prove that $[\![\mathsf{F}^{+}\langle\!\langle t^{o}\rangle\!\rangle]\!]_V \Longrightarrow [\![t]\!]_M = \mathit{tt}$, $\neg\, [\![\mathsf{F}^{-}\langle\!\langle t^{o}\rangle\!\rangle]\!]_V \Longrightarrow [\![t]\!]_M = \mathit{ff}$, and $[\![\mathsf{T}\langle\!\langle t\rangle\!\rangle]\!]_V \subseteq \lfloor [\![t]\!]_M \rfloor$ if $V(u) \subseteq \lfloor M(u) \rfloor$ for all free variables and nonstandard constants u occurring in t. Some of the cases deserve more justification:

- The $\mathsf{F}^{+}\langle\!\langle t \simeq u\rangle\!\rangle$ equation is sound because if the intersection of $\mathsf{T}\langle\!\langle t\rangle\!\rangle$ and $\mathsf{T}\langle\!\langle u\rangle\!\rangle$ is nonempty, then t and u must be equal (since they are singletons).
- The $\mathsf{F}^{-}\langle\!\langle t \simeq u\rangle\!\rangle$ equation is dual: If the union of $\mathsf{T}\langle\!\langle t\rangle\!\rangle$ and $\mathsf{T}\langle\!\langle u\rangle\!\rangle$ has more than one element, then t and u must be unequal.
- Universal quantification occurring positively can never yield true if the bound variable ranges over an approximated type. (In negative contexts, approximation compromises the encoding's completeness but not its soundness.)
- The *if then else* equation carefully distinguishes between the cases where the condition is *True*, *False*, and $\bot$. In the *True* case, it returns the *then* value; in the *False* case, it returns the *else* value; and in the $\bot$ case, it returns $\bot$ (none).
- The $\mathsf{T}\langle\!\langle c(t_1,\ldots,t_n)\rangle\!\rangle$ equation is as before. If any of the arguments t_j evaluates to none, the entire dot-join expression yields none.

Moreover, from a satisfying valuation V of the u_j's, we can construct a FOL model M such that $V(u_j) \subseteq \lfloor M(u_j) \rfloor$ for all u_j's, by defining $M(u_j)$ arbitrarily if $V(u_j) = \emptyset$ or at points where the partial function $V(u_j)$ is undefined. Hence, $[\![\mathsf{F}^{+}\langle\!\langle P\rangle\!\rangle]\!]_V$ implies $[\![P]\!]_M = \mathit{tt}$. $\square$

Although our translation is sound, a lot of precision is lost for $\simeq$ and $\forall$. Fortunately, by handling high-level definitional principles specially (as opposed to directly translating their FOL axiomatization), we can bypass the imprecise translation and increase the precision. This is covered in the next section.

5 Translation of Definitional Principles

5.1 Axiomatization of Simple Definitions

Once we extend the specification logic with simple definitions, we must also encode these in the FORL formula. More precisely, if c^{τ} is defined and an instance $c^{\tau'}$ occurs in a formula, we must conjoin c's definition with the formula, instantiating τ with τ'. This process must be repeated for any defined constants occurring in c's definition.

Given the command

$$\textsf{definition}\ c^{\tau}\ \textsf{where}\ c(\bar{x}) \simeq t$$

the naive approach would be to conjoin $\mathsf{F}^{+}\langle\!\langle \forall \bar{x}.\ c(\bar{x}) \simeq t \rangle\!\rangle$ with the FORL formula to satisfy and recursively do the same for any defined constants in t. However, there are two problems with this approach:

- If any of the variables $\bar{x}$ is of an approximated type, the equation $\mathsf{F}^{+}\langle\!\langle \forall \bar{x}.\ t \rangle\!\rangle = \mathsf{false}$ applies, and the axiom becomes unsatisfiable. This is sound but extremely imprecise, as it prevents the discovery of any model.
- Otherwise, the body of $\forall \bar{x}.\ c(\bar{x}) \simeq t$ is translated to $\mathsf{some}\ (\mathsf{T}\langle\!\langle c(\bar{x}) \rangle\!\rangle \cap \mathsf{T}\langle\!\langle t \rangle\!\rangle)$, which evaluates to false whenever $\mathsf{T}\langle\!\langle t \rangle\!\rangle$ is none for some values of $\bar{x}$.

Fortunately, we can take a shortcut and translate the definition directly to the following FORL axiom, bypassing F^{+} altogether (cf. Weber [31, p. 66]):

$$\forall x_1 \in \langle\!\langle \sigma_1 \rangle\!\rangle, \ldots, x_n \in \langle\!\langle \sigma_n \rangle\!\rangle \colon \mathsf{T}\langle\!\langle c(x_1, \ldots, x_n) \rangle\!\rangle \simeq \mathsf{T}\langle\!\langle t \rangle\!\rangle.$$

Theorem 5.1. *The encoding of Sect. 4.2 extended with simple definitions is sound.*

Proof. Any FORL valuation V that satisfies the FORL axiom for a constant c can be extended into a FOL model M that satisfies the corresponding FOL axiom, by setting $M(c)(\bar{v}) = [\![t]\!]_M(\bar{v})$ for any values $\bar{v}$ at which $V(c)$ is not defined (either because $\bar{v}$ is not representable in FORL or because the partial function $V(c)$ is not defined at that point). The apparent circularity in $M(c)(\bar{v}) = [\![t]\!]_M(\bar{v})$ is harmless, because simple definitions are required to be acyclic and so we can construct M one constant at a time. $\square$

5.2 Axiomatization of Algebraic Datatypes and Recursive Functions

The FORL axiomatization of algebraic datatypes follows the lines of Kuncak and Jackson [21]. Let $\kappa = C_1\ \textsf{of}\ (\sigma_{11}, \ldots, \sigma_{1n_1}) \mid \cdots \mid C_\ell\ \textsf{of}\ (\sigma_{\ell 1}, \ldots, \sigma_{\ell n_\ell})$ be a datatype instance. With each constructor C_i, we associate a discriminator $D_i^{\kappa \to o}$ and n selectors $S_{ik}^{\kappa \to \sigma_{ik}}$ obeying $D_j(C_i(\bar{x})) \simeq (i \simeq j)$ and $S_{ik}(C_i(x_1, ..., x_n)) \simeq x_k$. For example, the type $\alpha\ list$ is assigned the discriminators *nilp* and *consp* and the selectors *head* and *tail*:[4]

[4] These names were chosen for readability; any fresh names would do.

$$\begin{array}{ccc} \mathit{nilp}(\mathit{Nil}) \simeq \mathit{True} & \mathit{nilp}(\mathit{Cons}(x, xs)) \simeq \mathit{False} & \mathit{head}(\mathit{Cons}(x, xs)) \simeq x \\ \mathit{consp}(\mathit{Nil}) \simeq \mathit{False} & \mathit{consp}(\mathit{Cons}(x, xs)) \simeq \mathit{True} & \mathit{tail}(\mathit{Cons}(x, xs)) \simeq xs. \end{array}$$

The discriminator and selector view almost always results in a more efficient SAT encoding than the constructor view because it breaks high-arity constructors into several low-arity discriminators and selectors, declared as follows (for all possible i, k):

$$\textbf{var}\ \emptyset \subseteq D_i \subseteq \langle\!\langle \kappa \rangle\!\rangle \qquad\qquad \textbf{var}\ \emptyset \subseteq S_{ik} \subseteq \langle\!\langle \kappa \to \sigma_{ik} \rangle\!\rangle$$

The predicate D_i is directly coded as a set of atoms, rather than as a function to $\{\mathsf{a}_1, \mathsf{a}_2\}$.

Let $C_i\langle r_1, \ldots, r_n\rangle$ stand for $S_{i1}.r_1 \cap \cdots \cap S_{in}.r_n$ if $n \geq 1$, and $C_i\langle\rangle = D_i$ for parameterless constructors. Intuitively, $C_i\langle r_1, \ldots, r_n\rangle$ represents the constructor C_i with arguments $r_1, \ldots, r_n$ at the FORL level [10]. A faithful axiomatization of datatypes in terms of D_i and S_{ik} involves the following axioms (for all possible i, j, k):

$$\begin{array}{rl} \textsc{Disjoint}_{ij}\text{:} & \mathsf{no}\ D_i \cap D_j \quad \text{for } i < j \\ \textsc{Exhaustive}\text{:} & D_1 \cup \cdots \cup D_\ell \simeq \langle\!\langle \kappa \rangle\!\rangle \\ \textsc{Selector}_{ik}\text{:} & \forall y \in \langle\!\langle \kappa \rangle\!\rangle\text{: } \mathsf{if}\ y \subseteq D_i\ \mathsf{then\ one}\ y.S_{ik}\ \mathsf{else\ no}\ y.S_{ik} \\ \textsc{Unique}_i\text{:} & \forall x_1 \in \langle\!\langle \sigma_1 \rangle\!\rangle, \ldots, x_{n_i} \in \langle\!\langle \sigma_{n_i} \rangle\!\rangle\text{: } \mathsf{lone}\ C_i\langle x_1, \ldots, x_{n_i}\rangle \\ \textsc{Generator}_i\text{:} & \forall x_1 \in \langle\!\langle \sigma_1 \rangle\!\rangle, \ldots, x_{n_i} \in \langle\!\langle \sigma_{n_i} \rangle\!\rangle\text{: } \mathsf{some}\ C_i\langle x_1, \ldots, x_{n_i}\rangle \\ \textsc{Acyclic}\text{:} & \mathsf{no}\ \mathit{sup}_\kappa \cap \mathsf{iden}. \end{array}$$

In the last axiom, sup_κ denotes the proper superterm relation for κ. We will see shortly how to derive it from the selectors.

DISJOINT and EXHAUSTIVE ensure that the discriminators partition $\langle\!\langle \kappa \rangle\!\rangle$. The four remaining axioms, sometimes called the SUGA axioms (after the first letter of each axiom name), ensure that selectors are functions whose domain is given by the corresponding discriminator (SELECTOR), that constructors are total functions (UNIQUE and GENERATOR), and that datatype values cannot be proper superterms of themselves (ACYCLIC). The injectivity of constructors follows from the functionality of selectors.

With this axiomatization, occurrences of $C_i(u_1, \ldots, u_n)$ in FOL are simply mapped to $C_i\langle \mathsf{T}\langle\!\langle u_1 \rangle\!\rangle, \ldots, \mathsf{T}\langle\!\langle u_n \rangle\!\rangle\rangle$, whereas $\mathit{case}\ t\ \mathit{of}\ C_1(\bar{x}_1) \Rightarrow u_1 \mid \ldots \mid C_\ell(\bar{x}_\ell) \Rightarrow u_\ell$ is coded as

$$\mathsf{if}\ \mathsf{T}\langle\!\langle t \rangle\!\rangle \subseteq D_1\ \mathsf{then}\ \mathsf{T}\langle\!\langle u_1^\star \rangle\!\rangle\ \mathsf{else\ if} \ldots \mathsf{else\ if}\ \mathsf{T}\langle\!\langle t \rangle\!\rangle \subseteq D_\ell\ \mathsf{then}\ \mathsf{T}\langle\!\langle u_\ell^\star \rangle\!\rangle\ \mathsf{else\ none},$$

where $u_i^\star$ denotes the term u_i in which all occurrences of the variables $\bar{x}_i = x_{i1}, \ldots, x_{in_i}$ are replaced with the corresponding selector expressions $S_{i1}(t), \ldots, S_{in_i}(t)$.

Unfortunately, the SUGA axioms admit no finite models if the type κ is recursive (and hence infinite), because they force the existence of infinitely many values. The solution is to leave GENERATOR out, yielding SUA. The SUA axioms characterize precisely the subterm-closed finite substructures of an algebraic datatype. In a two-valued logic, this is generally unsound, but Kuncak and Jackson [21] showed that omitting GENERATOR is sound for *existential–bounded-universal* (EBU) sentences—namely, the formulas whose prenex normal forms contain no unbounded universal quantifiers ranging over datatypes.

In contrast, in our three-valued setting, omitting GENERATOR is always sound. The construct $C_i\langle r_1,\ldots,r_{n_i}\rangle$ sometimes returns none for non-none arguments, but this is not a problem since our translation of Sect. 4.2 is designed to cope with partiality. Non-EBU formulas such as $\mathit{True} \vee \forall n^{\mathit{nat}}.\; P(n)$ become analyzable when moving to a three-valued logic. This is especially important for complex specifications, because they are likely to contain non-EBU parts that are not needed for finding a model.

Example 5.1. The *nat list* instance of α *list* would be axiomatized as follows:

DISJOINT:	no $\mathit{nilp} \cap \mathit{consp}$
EXHAUSTIVE:	$\mathit{nilp} \cup \mathit{consp} \simeq \langle\!\langle \mathit{nat\ list} \rangle\!\rangle$
SELECTOR$_{\mathit{head}}$:	$\forall \mathit{ys} \in \langle\!\langle \mathit{nat\ list} \rangle\!\rangle$: if $\mathit{ys} \subseteq \mathit{consp}$ then one $\mathit{ys}.\mathit{head}$ else no $\mathit{ys}.\mathit{head}$
SELECTOR$_{\mathit{tail}}$:	$\forall \mathit{ys} \in \langle\!\langle \mathit{nat\ list} \rangle\!\rangle$: if $\mathit{ys} \subseteq \mathit{consp}$ then one $\mathit{ys}.\mathit{tail}$ else no $\mathit{ys}.\mathit{tail}$
UNIQUE$_{\mathit{Nil}}$:	lone $\mathit{Nil}\langle\rangle$
UNIQUE$_{\mathit{Cons}}$:	$\forall x \in \langle\!\langle \mathit{nat} \rangle\!\rangle, \mathit{xs} \in \langle\!\langle \mathit{nat\ list} \rangle\!\rangle$: lone $\mathit{Cons}\langle x, \mathit{xs}\rangle$
ACYCLIC:	no $\mathit{sup}_{\mathit{nat\ list}} \cap$ iden with $\mathit{sup}_{\mathit{nat\ list}} = \mathit{tail}^{+}$.

Examples of subterm-closed list substructures using traditional notation are $\{[], [0], [1]\}$ and $\{[], [1], [2,1], [0,2,1]\}$. In contrast, $L = \{[], [1,1]\}$ is not subterm-closed, because $\mathit{tail}([1,1]) = [1] \notin L$. Given a cardinality, Kodkod systematically enumerates all corresponding subterm-closed list substructures. ■

To generate the proper superterm relation needed for ACYCLIC, we must consider the general case of mutually recursive datatypes. We start by computing the datatype dependency graph, in which vertices are labeled with datatypes and arcs with selectors. For each selector $S^{\kappa\to\kappa'}$, we add an arc from κ to κ' labeled S. Next, we compute for each datatype a regular expression capturing the nontrivial paths in the graph from the datatype to itself. This can be done using Kleene's construction [18; 19, pp. 51–53]. The proper superterm relation is obtained from the regular expression by replacing concatenation with relational composition, alternative with set union, and repetition with transitive closure.

Example 5.2. Let *sym* be an atomic type. The definitions on the left-hand side give rise to the dependency graph on the right-hand side:

datatype $\alpha\ \mathit{list} = \mathit{Nil} \mid \mathit{Cons}$ **of** $(\alpha, \alpha\ \mathit{list})$
datatype $\mathit{tree} = \mathit{Leaf}$ **of** $\mathit{sym} \mid \mathit{Node}$ **of** $\mathit{tree\ list}$

$$\begin{tikzcd} \mathit{tree} \arrow[r, bend left, "\mathit{children}"] & \mathit{tree\ list} \arrow[l, bend left, "\mathit{head}"] \arrow[loop right, "\mathit{tail}"] \end{tikzcd}$$

The selector associated with *Node* is called *children*. The superterm relations are

$$\mathit{sup}_{\mathit{tree}} = (\mathit{children}.\mathit{tail}^{*}.\mathit{head})^{+} \qquad \mathit{sup}_{\mathit{tree\ list}} = (\mathit{tail} \cup \mathit{head}.\mathit{children})^{+}.$$

Notice that in the presence of polymorphism, instances of sequentially declared datatypes can be mutually recursive. ■

With a suitable axiomatization of datatypes as subterm-closed substructures, it is easy to encode **primrec** definitions. A recursive equation $f(C_i(x_1^{\sigma_1},\ldots,x_m^{\sigma_m}), z_1^{\sigma'_1},\ldots,z_n^{\sigma'_n}) \simeq t$ is translated to

$$\forall y \in D_i,\, z_1 \in \langle\!\langle \sigma_1' \rangle\!\rangle, \ldots, z_n \in \langle\!\langle \sigma_n' \rangle\!\rangle\colon\ \mathsf{T}\langle\!\langle f(y, z_1, \ldots, z_n) \rangle\!\rangle \simeq \mathsf{T}\langle\!\langle t^\star \rangle\!\rangle,$$

where $t^\star$ is obtained from t by replacing the variables x_i with the selector expressions $S_i(y)$. By quantifying over the constructed values y rather than on the arguments to the constructors, we reduce the number of copies of the quantified body by a factor of $|\langle\!\langle \sigma_1 \rangle\!\rangle| \cdot \ldots \cdot |\langle\!\langle \sigma_n \rangle\!\rangle| \,/\, |\langle\!\langle \kappa \rangle\!\rangle|$ in the SAT problem. Although we focus here on primitive recursion, general well-founded recursion with non-overlapping pattern matching (as defined using, say, Isabelle's **function** package [20]) can be handled in essentially the same way.

Example 5.3. The recursive function *cat* from Sect. 3.4 is translated to

$$\forall ys \in nilp,\, zs \in \langle\!\langle \alpha\ list \rangle\!\rangle\colon\ zs\cdot(ys\cdot cat) \simeq zs$$
$$\forall ys \in consp,\, zs \in \langle\!\langle \alpha\ list \rangle\!\rangle\colon\ zs\cdot(ys\cdot cat) \simeq Cons\langle ys\cdot head,\, zs\cdot((ys\cdot tail)\cdot cat)\rangle. \quad \blacksquare$$

Theorem 5.2. *The encoding of Sect. 5.1 extended with algebraic datatypes and primitive recursion is sound.*

Proof. Kuncak and Jackson [21] proved that SUA axioms precisely describe subterm-closed finite substructures of an algebraic datatype, and showed how to generalize this result to mutually recursive datatypes. This means that we can extend the valuation of the descriptors and selectors to obtain a model. For recursion, we can prove $[\![\mathsf{T}\langle\!\langle f(C(x_1, \ldots, x_m), z_1, \ldots, z_n) \rangle\!\rangle]\!]_V \subseteq \lfloor [\![f(C(x_1, \ldots, x_m), z_1, \ldots, z_n)]\!]_M \rfloor$ by structural induction on the value of the first argument to f and extend f's model as in the proof of Theorem 5.1, exploiting the injectivity of constructors. $\square$

5.3 Axiomatization of (Co)inductive Predicates

With datatypes and recursion in place, we are ready to consider (co)inductive predicates. Recall from Sect. 3.2 that an inductive predicate is the least fixed point p of the equation $p(\bar{x}) \simeq t[p]$ (where $t[p]$ is some formula involving p) and a coinductive predicate is the greatest fixed point. A first intuition would be to take $p(\bar{x}) \simeq t[p]$ as p's definition. In general, this is unsound since it underspecifies p, but there are two important cases for which this method is sound.

First, if the recursion in $p(\bar{x}) \simeq t[p]$ is well-founded, the equation admits exactly one solution [13]; we can safely use it as p's specification, and encode it the same way as a recursive function (Sect. 5.2). To ascertain wellfoundedness, we can perform a simple syntactic check to ensure that each recursive call peels off at least one constructor. Alternatively, we can invoke an off-the-shelf termination prover such as AProVE [11] or Isabelle's *lexicographic_order* tactic [6]. Given introduction rules of the form $p(\bar{t}_{i1}) \wedge \cdots \wedge p(\bar{t}_{i\ell_i}) \wedge Q_i \longrightarrow p(\bar{u}_i)$ for $i \in \{1, \ldots, n\}$, the prover attempts to exhibit a well-founded relation R such that $\bigwedge_{i=1}^{n} \bigwedge_{j=1}^{\ell_i} Q_i \longrightarrow \langle t_{ij}, u_i \rangle \in R$ holds. This is the approach implemented in Nitpick.

Second, if p is inductive and occurs negatively in the formula, we can replace these occurrences by a fresh constant q satisfying $q(\bar{x}) \simeq t[q]$. The resulting formula is equisatisfiable to the original formula: Since p is a least fixed point, q overapproximates p

and thus $\neg\, q(\bar{x}) \Longrightarrow \neg\, p(\bar{x})$. Dually, this method can also handle positive occurrences of coinductive predicates.

To deal with positive occurrences of inductive predicates, we adapt a technique from bounded model checking [3]: We replace these occurrences of p by a fresh predicate r_k defined by the FOL equations

$$r_0(\bar{x}) \simeq \mathit{False} \qquad\qquad r_{\mathit{Suc}\, n}(\bar{x}) \simeq t[r_n],$$

which corresponds to p unrolled k times. In essence, we have made the predicate well-founded by introducing a counter that decreases by one with each recursive call. The above equations are primitive recursive over the datatype *nat* and can be translated using the approach shown in Sect. 5.2. The unrolling comes at a price: The search space for r_k is k times that of p directly encoded as $p(\bar{x}) \simeq t[p]$.

The situation is mirrored for coinductive predicates: Negative occurrences are replaced by the overapproximation r_k defined by

$$r_0(\bar{x}) \simeq \mathit{True} \qquad\qquad r_{\mathit{Suc}\, n}(\bar{x}) \simeq t[r_n].$$

Theorem 5.3. *The encoding of Sect. 5.2 extended with (co)inductive predicates is sound.*

Proof. We consider only inductive predicates; coinduction is dual. If p is well-founded, the fixed point equation fully characterizes p [13], and the proof is identical to that of primitive recursion in Theorem 5.2 but with recursion induction instead of structural induction. If p is not well-founded, $q \simeq t[q]$ is satisfied by several q's, and by Knaster–Tarski $p \sqsubseteq q$. Substituting q for p's negative occurrences in the FORL formula strengthens it, which is sound. For the positive occurrences, we have $r_0 \sqsubseteq \cdots \sqsubseteq r_k \sqsubseteq p$ by monotonicity of the inductive definition; substituting r_k for p's positive occurrences strengthens the formula. □

Incidentally, we can mobilize FORL's transitive closure to avoid the explicit unrolling for an important class of inductive predicates, *linear inductive predicates*, whose introduction rules are of the form $Q \longrightarrow p(\bar{u})$ (the *base rules*) or $p(\bar{t}) \wedge Q \longrightarrow p(\bar{u})$ (the *step rules*). Informally, the idea is to replace positive occurrences of $p(\bar{x})$ with

$$\exists \bar{x}_0.\; p_{\text{base}}(\bar{x}_0) \wedge p^*_{\text{step}}(\bar{x}_0, \bar{x}),$$

where $p_{\text{base}}(\bar{x}_0)$ iff $p(\bar{x}_0)$ can be deduced from a base rule, $p_{\text{step}}(\bar{x}_0, \bar{x})$ iff $p(\bar{x})$ can be deduced by applying one step rule assuming $p(\bar{x}_0)$, and p^*_{step} is the reflexive transitive closure of p_{step}. For example, an inductive reachability predicate $\mathit{reach}(s)$ defined inductively would be coded as a set of initial states $\mathit{reach}_{\text{base}}$ and the small-step transition relation $\mathit{reach}_{\text{step}}$. The approach is not so different from explicit unrolling, since Kodkod internally unrolls the transitive closure to saturation. Nonetheless, on some problems the transitive closure approach is several times faster, presumably because Kodkod unfolds the relation inline instead of introducing an explicit counter.

5.4 Axiomatization of Coalgebraic Datatypes and Corecursive Functions

Coalgebraic datatypes are similar to algebraic datatypes, but they allow infinite values. For example, the infinite lists $[0, 0, \ldots]$ and $[0, 1, 2, 3, \ldots]$ are possible values of the type *nat llist* of coalgebraic (lazy) lists over natural numbers.

In principle, we could use the same SUA axiomatization for codatatypes as for datatypes (Sect. 5.2). This would exclude all infinite values but nonetheless be sound (although incomplete). However, in practice, infinite values often behave in surprising ways; excluding them would also exclude many interesting models.

It turns out we can modify the SUA axiomatization to support an important class of infinite values, namely those that are ω-regular. For lazy lists, this means lasso-shaped objects such as $[0,0,\ldots]$ and $[8,1,2,1,2,\ldots]$ (where the cycle $1,2$ is repeated infinitely).

The first step is to leave out the ACYCLIC axiom. However, doing only this is unsound, because we might obtain several atoms encoding the same value; for example, $\mathsf{a}_1 = \mathit{LCons}(0, \mathsf{a}_1)$, $\mathsf{a}_2 = \mathit{LCons}(0, \mathsf{a}_3)$, and $\mathsf{a}_3 = \mathit{LCons}(0, \mathsf{a}_2)$ all encode the infinite list $[0,0,\ldots]$. This violates the bisimilarity principle, according to which two values are equal unless they lead to different observations (the observations being $0,0,\ldots$).

For lazy lists, we add the definition

$$\begin{array}{l}\textbf{coinductive} \sim^{(\alpha\ \mathit{llist},\, \alpha\ \mathit{llist}) \to o} \textbf{ where}\\ \mathit{LNil} \sim \mathit{LNil}\\ x \simeq x' \wedge xs \sim xs' \longrightarrow \mathit{LCons}(x, xs) \sim \mathit{LCons}(x', xs')\end{array}$$

and we require that $\simeq$ coincides with $\sim$ on $\alpha\ \mathit{llist}$ values. More generally, we generate mutual coinductive definitions of $\sim$ for all the codatatypes. For each constructor $C^{(\sigma_1,\ldots,\sigma_n)\to\sigma}$, we add an introduction rule

$$x_1 \approx_1 x_1' \wedge \cdots \wedge x_n \approx_n x_n' \longrightarrow C(x_1,\ldots,x_n) \sim C(x_1',\ldots,x_n'),$$

where $\approx_i$ is $\sim^{(\sigma_i,\sigma_i)\to o}$ if σ_i is a codatatype and $\simeq$ otherwise. Finally, for each codatatype κ, we add the axiom

$$\text{BISIMILAR:}\quad \forall y, y' \in \langle\!\langle \kappa \rangle\!\rangle\colon\ y \sim y' \longrightarrow y \simeq y'.$$

With the SUB (SU plus BISIMILAR) axiomatization in place, it is easy to encode **coprimrec** definitions. A corecursive equation $f(y_1^{\sigma_1},\ldots,y_n^{\sigma_1}) \simeq t$ is translated to

$$\forall y_1 \in \langle\!\langle \sigma_1 \rangle\!\rangle, \ldots, y_n \in \langle\!\langle \sigma_n \rangle\!\rangle\colon\ \mathsf{T}\langle\!\langle f(y_1,\ldots,y_n) \rangle\!\rangle \simeq \mathsf{T}\langle\!\langle t \rangle\!\rangle.$$

Theorem 5.4. *The encoding of Sect. 5.3 extended with coalgebraic datatypes and primitive corecursion is sound.*

Proof. Codatatypes correspond to final coalgebras. They are characterized by selectors, which are axiomatized by the SU axioms, and by finality, which is equivalent to the bisimilarity principle [16,26]. Our finite axiomatization gives a subterm-closed substructure of the coalgebraic datatype, which can be extended to yield a FOL model of the complete codatatype, as we did for algebraic datatypes in the proof of Theorem 5.2.

The soundness of the encoding of primitive corecursion is proved by coinduction. Given the equation $f(\bar{y}) \simeq t$, assuming that for each corecursive call $f(\bar{x})$ we have $[\![\mathsf{T}\langle\!\langle f(\bar{x}) \rangle\!\rangle]\!]_V \subseteq \lfloor [\![f(\bar{x})]\!]_M \rfloor$, we must show that $[\![\mathsf{T}\langle\!\langle f(\bar{y}) \rangle\!\rangle]\!]_V \subseteq \lfloor [\![f(\bar{y})]\!]_M \rfloor$. This follows from the soundness of the encoding of the constructs occurring in the right-hand side t and from the hypotheses. $\square$

6 Case Study: Lazy Lists

The codatatype α *llist* of lazy lists [26] is generated by the constructors $LNil^{\alpha\ llist}$ and $LCons^{\alpha\to\alpha\ llist\to\alpha\ llist}$. It is of particular interest to (counter)model finding because many basic properties of finite lists do not carry over to infinite lists, often in baffling ways. To illustrate this, we conjecture that appending *ys* to *xs* yields *xs* iff *ys* is *LNil*:

$$(lcat(xs, ys) \simeq xs) \simeq (ys \simeq LNil).$$

The function *lcat* is defined corecursively in Sect. 3.4. For the conjecture, our tool Nitpick immediately finds the countermodel $xs = ys = [0, 0, \ldots]$, in which a cardinality of 1 is sufficient for α and α *llist*, and the bisimilarity predicate $\sim$ is unrolled only once. Indeed, appending $[0, 0, \ldots] \neq []$ to $[0, 0, \ldots]$ leaves $[0, 0, \ldots]$ unchanged.

The next example requires the following lexicographic order predicate:

$$\begin{array}{l}
\textsf{coinductive}\ \preceq^{(nat\ llist,\, nat\ llist)\to o}\ \textsf{where} \\
LNil \preceq xs \\
x \le y \longrightarrow LCons(x, xs) \preceq LCons(y, ys) \\
xs \preceq ys \longrightarrow LCons(x, xs) \preceq LCons(x, ys)
\end{array}$$

The intention of this definition is to define a linear order on lazy lists of natural numbers, and hence the following properties should hold:

$$\begin{array}{rlrl}
\text{REFL:} & xs \preceq xs & \text{ANTISYM:} & xs \preceq ys \land ys \preceq xs \longrightarrow xs \simeq ys \\
\text{LINEAR:} & xs \preceq ys \lor ys \preceq xs & \text{TRANS:} & xs \preceq ys \land ys \preceq zs \longrightarrow xs \preceq zs.
\end{array}$$

However, Nitpick finds a counterexample for ANTISYM: $xs = [1, 1]$ and $ys = [1]$. On closer inspection, the assumption $x \le y$ of the second introduction rule for $\preceq$ should have been $x < y$; otherwise, any two lists *xs*, *ys* with the same head satisfy $xs \preceq ys$. Once we repair the specification, no more counterexamples are found for the four properties up to cardinality 6 for *nat* and *nat llist*, within Nitpick's default time limit of 30 seconds. This is a strong indication that the properties hold. Andreas Lochbihler used Isabelle to prove all four properties [23].

7 Related Work

The encoding of algebraic datatypes in FORL has been studied by Kuncak and Jackson [21] and Dunets et al. [10]. Kuncak and Jackson focused on lists and trees. Dunets et al. showed how to handle primitive recursion; their approach to recursion is similar to ours, but the use of a two-valued logic compelled them to generate additional definedness guards. The unrolling of inductive predicates was inspired by bounded model checking [3] and by the Alloy idiom for state transition systems [15, pp. 172–175].

Another inspiration has been Weber's higher-order model finder Refute [31]. It uses a three-valued logic, but sacrifices soundness for precision. Datatypes are approximated by subterm-closed substructures [31, pp. 58–64] that contain *all* datatype values built using up to *k* nested constructors. This scheme proved disadvantageous in practice, because it generally requires higher cardinalities to obtain the same models as with Kuncak and Jackson's approach. Weber handled (co)inductive predicates by expanding their HOL definition, which in practice does not scale beyond a cardinality of 3 for the predicate's domain because of the higher-order quantifier.

The Nitpick tool, which implements the techniques presented here, is described in a separate paper [5] that covers the handling of higher-order quantification and functions. The paper also presents an evaluation of the tool on various Isabelle/HOL theories, where it competes against Quickcheck [2] and Refute [31], as well as two case studies.

8 Conclusion

Despite recent advances in lightweight formal methods, there remains a wide gap between specification languages that lend themselves to automatic analysis and those that are used in actual formalizations. As an example, infinite types are ubiquitous, yet most model finders either spin forever [9, 24], give up immediately [8], or are unsound [1; 28, p. 164; 31] on finitely unsatisfiable formulas.

We identified several commonly used definitional principles and showed how to encode them in first-order relational logic (FORL), the logic supported by the Kodkod model finder and the Alloy Analyzer. Our main contribution has been to develop three ways to translate (co)inductive predicates to FORL, based on wellfoundedness, polarity, and linearity. Other contributions have been to formulate an axiomatization of coalgebraic datatypes that caters for infinite (ω-regular) values and to devise a procedure that computes the acyclicity axiom for mutually recursive datatypes.

Our experience with the counterexample generator Nitpick has shown that the techniques scale to handle real-world specifications, including a security type system and a hotel key card system [5]. Although the tool is fairly new, one user has already reported saving several hours of failed proof attempts thanks to its support for codatatypes and coinductive predicates while developing a formal theory of infinite process traces [22].

Acknowledgment. I want to thank Sascha Böhme, Lukas Bulwahn, Andreas Lochbihler, Tobias Nipkow, Mark Summerfield, and the anonymous reviewers for suggesting many improvements to this paper, and Alexander Krauss for helping to structure it.

References

1. W. Ahrendt. Deductive search for errors in free data type specifications using model generation. In A. Voronkov, ed., *CADE-18*, vol. 2392 of *LNAI*, pp. 211–225. Springer, 2002.
2. S. Berghofer and T. Nipkow. Random testing in Isabelle/HOL. In J. Cuellar and Z. Liu, eds., *SEFM 2004*, pp. 230–239. IEEE C.S., 2004.
3. A. Biere, A. Cimatti, E. M. Clarke, and Y. Zhu. Symbolic model checking without BDDs. In R. Cleaveland, ed., *TACAS '99*, vol. 1579 of *LNCS*, pp. 193–207. Springer, 1999.
4. J. C. Blanchette and A. Krauss. Monotonicity inference for higher-order formulas. In J. Giesl and R. Hähnle, eds., *IJCAR 2010*, LNCS. Springer, 2010. To appear.
5. J. C. Blanchette and T. Nipkow. Nitpick: A counterexample generator for higher-order logic based on a relational model finder. In M. Kaufmann and L. Paulson, eds., *ITP-10*, LNCS. Springer, 2010. To appear.

6. L. Bulwahn, A. Krauss, and T. Nipkow. Finding lexicographic orders for termination proofs in Isabelle/HOL. In K. Schneider and J. Brandt, eds., *TPHOLs 2007*, vol. 4732 of *LNCS*, pp. 38–53. Springer, 2007.
7. A. Church. A formulation of the simple theory of types. *J. Symb. Log.*, 5:56–68, 1940.
8. K. Claessen and A. Lillieström. Automated inference of finite unsatisfiability. In R. A. Schmidt, ed., *CADE-22*, vol. 5663 of *LNAI*, pp. 388–403. Springer, 2009.
9. K. Claessen and N. Sörensson. New techniques that improve MACE-style model finding. In *MODEL*, 2003.
10. A. Dunets, G. Schellhorn, and W. Reif. Bounded relational analysis of free datatypes. In B. Beckert and R. Hähnle, eds., *TAP 2008*, vol. 4966 of *LNCS*, pp. 99–115. Springer, 2008.
11. J. Giesl, P. Schneider-Kamp, and R. Thiemann. AProVE 1.2: Automatic termination proofs in the dependency pair framework. In *IJCAR 2006*, vol. 4130 of *LNAI*, pp. 281–286, 2006.
12. M. J. C. Gordon and T. F. Melham, eds. *Introduction to HOL: A Theorem Proving Environment for Higher Order Logic*. Cambridge University Press, 1993.
13. J. Harrison. Inductive definitions: Automation and application. In E. T. Schubert, P. J. Windley, and J. Alves-Foss, eds., *TPHOLs 1995*, vol. 971 of *LNCS*, pp. 200–213. Springer, 1995.
14. J. Harrison. HOL Light: A tutorial introduction. In *FMCAD '96*, vol. 1166 of *LNCS*, pp. 265–269. Springer, 1996.
15. D. Jackson. *Software Abstractions: Logic, Language, and Analysis*. MIT Press, 2006.
16. B. Jacobs and J. Rutten. A tutorial on (co)algebras and (co)induction. *Bull. EATCS*, 62:222–259, 1997.
17. S. C. Kleene. On notation for ordinal numbers. *J. Symb. Log.*, 3(4):150–155, 1938.
18. S. C. Kleene. Representation of events in nerve nets and finite automata. In J. McCarthy and C. Shannon, eds., *Automata Studies*, pp. 3–42. Princeton University Press, 1956.
19. D. C. Kozen. *Automata and Computability*. Undergrad. Texts in C.S. Springer, 1997.
20. A. Krauss. Partial and nested recursive function definitions in higher-order logic. *J. Auto. Reas.*, 44(4):303–336, 2009.
21. V. Kuncak and D. Jackson. Relational analysis of algebraic datatypes. In H. C. Gall, ed., *ESEC/FSE 2005*, 2005.
22. A. Lochbihler. Private communication, 2009.
23. A. Lochbihler. Coinduction. In G. Klein, T. Nipkow, and L. C. Paulson, eds., *The Archive of Formal Proofs*. `http://afp.sourceforge.net/entries/Coinductive.shtml`, Feb. 2010.
24. W. McCune. A Davis–Putnam program and its application to finite first-order model search: Quasigroup existence problems. Technical report, ANL, 1994.
25. T. Nipkow, L. C. Paulson, and M. Wenzel. *Isabelle/HOL: A Proof Assistant for Higher-Order Logic*, vol. 2283 of *LNCS*. Springer, 2002.
26. L. C. Paulson. A fixedpoint approach to implementing (co)inductive definitions. In A. Bundy, ed., *CADE-12*, vol. 814 of *LNAI*, pp. 148–161. Springer, 1994.
27. T. Ramananandro. Mondex, an electronic purse: Specification and refinement checks with the Alloy model-finding method. *Formal Asp. Comput.*, 20(1):21–39, 2008.
28. J. M. Schumann. *Automated Theorem Proving in Software Engineering*. Springer, 2001.
29. K. Slind and M. Norrish. A brief overview of HOL4. In O. A. Mohamed, C. M. noz, and S. Tahar, eds., *TPHOLs 2008*, vol. 5170 of *LNCS*, pp. 28–32, 2008.
30. E. Torlak and D. Jackson. Kodkod: A relational model finder. In O. Grumberg and M. Huth, eds., *TACAS 2007*, vol. 4424 of *LNCS*, pp. 632–647. Springer, 2007.
31. T. Weber. *SAT-Based Finite Model Generation for Higher-Order Logic*. Ph.D. thesis, Dept. of Informatics, T.U. München, 2008.
32. M. Wenzel. Type classes and overloading in higher-order logic. In E. L. Gunter and A. Felty, eds., *TPHOLs 1997*, vol. 1275 of *LNCS*, pp. 307–322. Springer, 1997.