Observational Ultraproducts of Polynomial Coalgebras*

Robert Goldblatt
Centre for Logic, Language and Computation
Victoria University of Wellington, New Zealand
Rob.Goldblatt@vuw.ac.nz

20 December 2002

Contents

Abstract

Coalgebras of polynomial functors constructed from sets of observable elements have been found useful in modelling various kinds of data types and state-transition systems. This paper continues the study of equational logic and model theory for polynomial coalgebras begun in [10], where it was shown that Boolean combinations of equations between terms of observable type form a natural language of *observable formulas* for specifying properties of polynomial coalgebras, and for giving a Hennessy-Milner style logical characterisation of observational indistinguishability (bisimilarity) of states.

Here we give a structural characterisation of classes of coalgebras definable by observable formulas. This is an analogue for polynomial coalgebras of Birkhoff's celebrated characterisation of equationally definable classes of abstract algebras as being those closed under homomorphic images, subalgebras, and direct products. The coalgebraic characterisation involves new notions of *observational ultraproduct* and *ultrapower* of coalgebras, obtained from the classical construction of ultraproducts by

*Prepared using Paul Taylor's `diagrams` package.

deleting states that assign "nonstandard" values to terms of observable type. A class of polynomial coalgebras is shown to be the class of all models of a set of observable formulas if, and only if, it is closed under images of bisimulation relations, disjoint unions and observational ultrapowers.

Observational ultraproducts are also used to discuss compactness—which holds only under limited conditions; to characterize finitely axiomatizable classes of polynomial coalgebras; and to show that there are axiomatizable classes that are not finitely axiomatizable.

1 Introduction and Overview

Let $T : \mathbf{Set} \to \mathbf{Set}$ be a functor on the category of sets. A *T-coalgebra* is a pair (A, α) with A being a set and α a function of the form $A \to TA$. This notion has proven useful in modelling data structures such as lists, streams and trees; transition systems such as automata; and classes in object-oriented programming languages [32, 23]. Generically A is viewed as a set of *states*, and α as a *transition structure*. The usefulness of the notion has fueled the development of a theory of "universal coalgebra" [34, 36, 13], by analogy with, and categorically dual to, the study of abstract algebras.

The present paper is a continuation of a study begun in [10] of equational logic and model theory for certain *polynomial* coalgebras. These are ones for which T is a *polynomial functor*, i.e. is constructed from constant-valued functors and the identity functor by forming products, exponential functors with constant exponent (which we will call *power* functors), and coproducts. Many of the examples of data structures cited in the previous paragraph arise from polynomial functors. Such functors are typically constructed from some fixed sets given in advance, with the members of these sets being thought of as "observable" elements. Computationally, the states of a coalgebra are regarded as not being directly accessible, but can only be investigated by performing certain "experiments" in the form of coalgebraic operations that yield observable values. Two states are *bisimilar* if no such experiment can distinguish them. This notion of bisimilarity, as meaning observational indistinguishability, is one of the key foundational concepts of process algebra. It derives from a notion of *bisimulation* introduced in [30] as a relation of mutual simulation between states of finite automata. Bisimilarity is the largest bisimulation relation between the state-sets of two coalgebras, and can be characterised in several ways, including set-theoretically using state-transitions and categorically using morphisms between coalgebras (see [10, Section 1] for references).

A formal language for expressing properties of polynomial coalgebras was introduced in [10] and its semantics defined. The core of the language is a calculus of *terms* for operations on coalgebras, based on a simple type theory that reflects the formation of polynomial functors. A distinctive feature of the terms is that they may have a single state-valued *parameter*, in addition to state-valued variables. The state variables can be bound by certain operations on terms (lambda-abstraction, case-formation), but the state parameter is not subject to binding. The symbol s is used for this parameter, which may be viewed as denoting the "current" state. A special role is played by terms that take observable values, since an "experiment" amounts to the evaluation of such a term. Theorem 7.2 of [10] established that two coalgebraic states are bisimilar precisely when they assign the same values to all ground (i.e. closed) terms of observable type. Moreover it was shown that Boolean combinations of equations between observable terms form a natural language of *observable formulas* for specifying coalgebraic properties. Two states are bisimilar precisely when they satisfy the same "rigid" observable formulas (rigidity essentially means that all state variables are bound).

The idea of characterising observationally equivalent states as being those satisfying the same formulas of some logical language was introduced by Hennessy and Milner [17]. They provided a simple propositional modal language that achieved this for certain labelled transition systems. Our results in [10] show that observable formulas provide the correspondingly appropriate language for polynomial coalgebras, and motivate further study of the metatheory of these formulas. In particular we consider the extent to which their role is analogous to the role played by *equations* in the general theory of universal algebras.

Now a cornerstone of classical equational logic is the "variety theorem" of Birkhoff [7], stating that a class of algebras is the class of all models of some set of equations iff it is closed under homomorphic images, subalgebras, and direct products. This paper aims to prove an analogous result for polynomial coalgebras, giving a structural characterisation of classes of coalgebras definable by observable formulas. For this purpose a new construction of *observational ultrapowers* is introduced. An ultrafilter U on a set I associates with each set B the standard ultrapower B^U, which is a quotient of the I-th power B^I of B, and which may be thought of as an "enlargement" obtained by adding new "nonstandard" elements to B. Any function $\theta : B \to C$ has a lifting to a function $\theta^U : B^U \to C^U$. In particular a T-coalgebra $\alpha : A \to TA$ lifts to $\alpha^U : A^U \to (TA)^U$. But α^U is not a T-coalgebra on A^U, since the latter would be a function of the form $A^U \to T(A^U)$. To overcome this problem we reduce A^U to a subset A^+ for which a suitable function $\alpha^+ : A^+ \to T(A^+)$ can be built using α and U.

The definition of A^+ can be understood by considering a ground (closed) term M whose type is a set D of observable values. The denotation of M in α is a function $[\![M]\!]_\alpha : A \to D$ giving M a value in each state of A. This lifts to a function $[\![M]\!]^U : A^U \to D^U$ which may assign some members x of A^U a nonstandard value in $D^U - D$. If however $[\![M]\!]^U(x) \in D$, and this property holds for every term M of observable type, then we call x an *observable state* of A^U. A^+ is the set of such observable states.

The construction of α^+ is intricate and involves a detailed analysis of the formation of T as a polynomial functor (see Theorem 4.1). α^+ validates the same observable formulas as α (Corollary 5.3). To prove this we have to establish a result (Theorem 5.2) about satisfaction of formulas by elements of α^+ which is the analogue in this setting of Łoś's Theorem in the standard theory of ultrapowers.

Our analogue of Birkhoff's theorem (Theorem 7.1) states that

> *a class of polynomial coalgebras is the class Mod Φ of all models of a set Φ of rigid observable formulas if, and only if, it is closed under images of bisimulation relations, disjoint unions and observational ultrapowers.*

The proof makes use of a certain *saturation* property of α^+ that holds for suitably chosen ultrafilters U. Saturation is used to show that if every ground observable formula valid in α is valid also in coalgebra β, then α^+ has enough states to guarantee that each state of β is bisimilar to a state of α^+, and so the bisimilarity relation from α^+ to β is surjective (Theorem 6.1). From this it follows (see Theorem 6.2) that

> *two coalgebras α and β are logically indistinguishable, in the sense that they validate the same ground observable formulas, if, and only if, they have observational ultrapowers α^+ and β^+ that are totally bisimilar and hence are observationally indistinguishable.*

The definition of α^+ is a special case of the more general construction of an observational ultra*product* $\Pi_U \alpha_i^+$ of a family $\langle \alpha_i : i \in I \rangle$ of coalgebras. This construction and the associated version of Łoś's Theorem is developed in full, and is used for a number of applications. These include:

(i) exhibiting partial failures of Łoś's Theorem (after Theorem 5.2);

(ii) showing that limited versions of the Compactness Theorem hold (Theorem 5.4);

(iii) giving conditions under which a class of coalgebras is equal to *Mod* φ for a single formula φ (Theorem 7.2); and

(iv) giving an example of a class of the form *Mod* Φ for some infinite Φ that is not finitely axiomatizable, hence not equal to *Mod* φ for any formula φ (end of Section 7).

Section 8 discusses the sense in which our main result is an analogue of the classical Variety Theorem, and also surveys and compares a number of contributions to the literature that deal with category-theoretic formulations and dualizations of Birkhoff's concepts.

$$
\begin{array}{l}
\mathit{size}(\mathsf{s}) = \ \mathsf{case}\ \mathit{children}(\mathsf{s})\ \mathsf{of} \\
\qquad \iota_1 u \mapsto \ \iota_2 1 \\
\qquad \iota_2 v \mapsto \ \mathsf{case}\ \mathit{size}(\pi_1 v)\ \mathsf{of} \\
\qquad\qquad \iota_1 u \mapsto \ \iota_1 * \\
\qquad\qquad \iota_2 n \mapsto \ \mathsf{case}\ \mathit{size}(\pi_2 v)\ \mathsf{of} \\
\qquad\qquad\qquad \iota_1 u \mapsto \iota_1 * \\
\qquad\qquad\qquad \iota_2 k \mapsto \iota_2(n + k + 1) \\
\qquad\qquad\qquad \mathsf{endcase} \\
\qquad\qquad \mathsf{endcase} \\
\qquad \mathsf{endcase}
\end{array}
$$

Figure 1: case terms

case Terms

The syntax of terms we use has the standard constructions from type theory, including pairing and projection terms for products of types, and λ-abstractions and application terms for power types (function spaces). Possibly less familiar is the "case" operation used to introduce terms associated with coproducts, and so we provide some motivation for it now.

The coproduct $A_1 + A_2$ of sets A_1 and A_2 is their disjoint union, and comes equipped with injective *insertion* functions $\iota_j : A_j \to A_1 + A_2$ for $j = 1$ and 2. Each element of $A_1 + A_2$ is equal to $\iota_j(a)$ for a unique j and a unique $a \in A_j$. Our syntax generates terms of the form

$$\mathsf{case}\ N\ \mathsf{of}\ [\iota_1 v_1 \mapsto M_1 \mid \iota_2 v_2 \mapsto M_2],$$

where N is a term taking values in $A_1 + A_2$, M_1 and M_2 take values in some other set B, and the v_j's are variables that take values in A_j *and are bound in the overall* case *term*. We will sometimes use the abbreviated form $\mathsf{case}(N, M_1, M_2)$ for this term. It is evaluated by first obtaining the value d of N in $A_1 + A_2$ and then, if d is equal to $\iota_j(a)$, evaluating M_j with v_j assigned value a, giving an element of B as the desired value.

Here is an illustration from [10], adapted from [25, Section 4], of the use of case-formation in coalgebraic specification. Let A be a set of (possibly infinite) binary trees. Each tree x either is a single node with no children, or has exactly two children obtained by deleting the top node of x. This gives an operation

$$\mathit{children} : A \longrightarrow 1 + (A \times A),$$

where $1 = \{*\}$; $\mathit{children}(x) = \iota_1 *$ when x has no children, and $\mathit{children}(x) = \iota_2(x_1, x_2)$ when x_1 and x_2 are the left and right children of x. There is a size (number of nodes) operation

$$\mathit{size} : A \longrightarrow 1 + \mathbb{N},$$

where $\mathbb{N}$ is the set of positive integers and $\mathit{size}(x) = \iota_1 *$ when x is infinite. The two operations can be "paired" into a single function

$$A \xrightarrow{\ \alpha\ } (1 + (A \times A)) \times (1 + \mathbb{N})$$

which is a coalgebra for the functor $T(X) = (1 + (X \times X)) \times (1 + \mathbb{N})$. The operations can be recovered from α as $\mathit{children} = \pi_1 \circ \alpha$ and $\mathit{size} = \pi_2 \circ \alpha$, where π_1 and π_2 are the left and right projections.

Now the size of a tree is 1 if it has no children, is infinite if at least one child is infinite, and otherwise is the sum of the sizes of the children plus 1. Thus our example *validates* the equation of Figure 1, in which the right-hand term is obtained by iteration of case-formation. (The word "endcase" used to mark the end of a term will not appear in our formal syntax.) Validity means that the equation is satisfied no matter what member of A is denoted by the state parameter s.

2 Review of Fundamentals

To make the paper reasonably self-contained, in this section we review the syntax and semantics of types, terms and formulas for polynomial coalgebras, as set out in [10], and describe the resulting logical characterisation of bisimulation.

2.1 Polynomial Functors and Coalgebras

Conventional notation for products, powers and coproducts of sets will be used. For $j = 1$ and 2, $\pi_j : A_1 \times A_2 \to A_j$ is the *projection* function from the *product* set $A_1 \times A_2$ onto A_j, i.e. $\pi_j(a_1, a_2) = a_j$. The *pairing* of two functions of the form $f_1 : A \to B_1$ and $f_2 : A \to B_2$ is the function $\langle f_1, f_2\rangle : A \to B_1 \times B_2$ given by $f(a) = (f_1(a), f_2(a))$. The *product* of two functions of the form $f_1 : A_1 \to B_1$ and $f_2 : A_2 \to B_2$ is the function $f_1 \times f_2 : A_1 \times A_2 \to B_1 \times B_2$ that maps (a_1, a_2) to $(f_1(a_1), f_2(a_2))$. Thus $\pi_j((f_1 \times f_2)(x)) = f_j(\pi_j(x))$.

As already mentioned, the coproduct $A_1 + A_2$ of sets A_1, A_2 is their disjoint union, with injective *insertion* functions $\iota_j : A_j \to A_1 + A_2$ for $j = 1$ and 2. Each element of $A_1 + A_2$ is equal to $\iota_j(a)$ for a unique j and a unique $a \in A_j$. The *coproduct* of two functions of the form $f_1 : A_1 \to B_1$ and $f_2 : A_2 \to B_2$ is the function $f_1 + f_2 : A_1 + A_2 \to B_1 + B_2$ that maps $\iota_j(a)$ to $\iota_j(f_j(a))$.

The *D-th power* of set A is the set A^D of all functions from set D to A. The D-th power of a function $f : A \to B$ is the function $f^D : A^D \to B^D$ having $f^D(g) = f \circ g$ for all $g : D \to A$. The *evaluation* function $eval : A^D \times D \to A$ has $eval(f, d) = f(d)$. For each $d \in D$ there is the *evaluation-at-d* function $ev_d : A^D \to A$ having $ev_d(f) = eval(f, d) = f(d)$.

The symbol $\circ\!\!\longrightarrow$ will be used for partial functions. Thus $f : A \circ\!\!\longrightarrow B$ means that f is a function with codomain B and domain $\mathrm{Dom}\, f \subseteq A$. We sometimes write $f(x)\!\downarrow$ to mean that $f(x)$ is defined, i.e. $x \in \mathrm{Dom}\, f$. Associated with each insertion $\iota_j : A_j \to A_1 + A_2$ is its partial inverse, the *extraction* function $\varepsilon_j : A_1 + A_2 \circ\!\!\longrightarrow A_j$ having $\varepsilon_j(y) = x$ iff $\iota_j(x) = y$. Thus $\mathrm{Dom}\, \varepsilon_j = \iota_j A_j$, i.e. $y \in \mathrm{Dom}\, \varepsilon_j$ iff $y = \iota_j(x)$ for some $x \in A_j$. These extraction functions play a vital role in the analysis of coalgebras built out of coproducts. Observe that the coproduct $f_1 + f_2$ of two functions has $(f_1 + f_2)(x) = \iota_j(f_j(\varepsilon_j(x)))$ for some j.

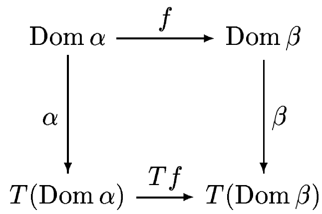
The identity function on a set A is denoted $\mathrm{id}_A : A \to A$. If A is a subset of B, then $A \hookrightarrow B$ is the inclusion function from A to B.

Polynomial functors are formed from the following constructions of endofunctors T : **Set** $\to$ **Set**.

- For a fixed set $D \neq \emptyset$, the *constant functor* $\bar{D}$ has $\bar{D}(A) = D$ on sets A and $\bar{D}(f) = \mathrm{id}_D$ on functions f.
- The *identity functor* Id has $\mathrm{Id}A = A$ and $\mathrm{Id}f = f$.
- The product $T_1 \times T_2$ of two functors has $T_1 \times T_2(A) = T_1A \times T_2A$, and, for a function $f : A \to B$, has $T_1 \times T_2(f)$ being the product function
$$T_1(f) \times T_2(f) : T_1A \times T_2A \to T_1B \times T_2B.$$
- The coproduct $T_1 + T_2$ of two functors has $T_1 + T_2(A) = T_1A + T_2A$, and for $f : A \to B$, has $T_1 + T_2(f)$ being the coproduct function
$$T_1(f) + T_2(f) : T_1A + T_2A \to T_1B + T_2B.$$
- The D-th power functor T^D of a functor T has $T^DA = (TA)^D$, and for $f : A \to B$, has $T^D(f)$ being the function $(T(f))^D : (TA)^D \to (TB)^D$ that acts by $g \mapsto T(f) \circ g$. Thus $T^D(f)(g)(d) = T(f)(g(d))$.

A functor T is *polynomial* if it is constructed from constant functors and Id by finitely many applications of products, coproducts and powers. Any functor formed as part of the construction of T is a *component* of T. A polynomial functor that does not have Id as a component must be constant.

A T-*coalgebra* is a pair (A, α) comprising a set A and a function $A \xrightarrow{\alpha} TA$. A is the set of *states* and α is the *transition structure* of the coalgebra. Note that A is determined as the domain $\mathrm{Dom}\,\alpha$ of α, so we can identify the coalgebra with its transition structure, i.e. a T-coalgebra is any function of the form $\alpha : \mathrm{Dom}\,\alpha \to T(\mathrm{Dom}\,\alpha)$. A *morphism* from T-coalgebra α to T-coalgebra β is a function $f : \mathrm{Dom}\,\alpha \to \mathrm{Dom}\,\beta$ between their state sets which commutes with their transition structures in the sense that $\beta \circ f = Tf \circ \alpha$, i.e. the following diagram commutes:



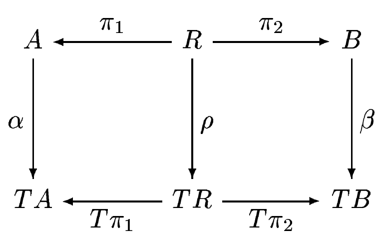
If $\mathrm{Dom}\,\alpha \subseteq \mathrm{Dom}\,\beta$, then α is a *subcoalgebra* of β iff the inclusion function $\mathrm{Dom}\,\alpha \hookrightarrow \mathrm{Dom}\,\beta$ is a morphism from α to β.

Every set $\{\alpha_i : i \in I\}$ of T-coalgebras has a *disjoint union* $\sum_I \alpha_i$, which is a T-coalgebra whose domain is the disjoint union of the $\mathrm{Dom}\,\alpha_i$'s and whose transition structure acts as α_j on the summand $\iota_j \mathrm{Dom}\,\alpha_j$ of $\mathrm{Dom} \sum_I \alpha_i$. More precisely, this transition is given by $\iota_j(a) \mapsto T(\iota_j)(\alpha_j(a))$, with the insertion $\iota_j : \mathrm{Dom}\,\alpha_j \to \mathrm{Dom} \sum_I \alpha_i$ being an injective morphism making α_j isomorphic to a subcoalgebra of the disjoint union (see [36, Section 4]).

In classical algebra and model theory it is conventional to assume that an algebra or model is based on a non-empty set. By contrast, here we always have the *empty T-coalgebra* $\emptyset \to T\emptyset$ with the empty transition structure. This forms an *initial object* in the category of T-coalgebras: there is a unique (empty) morphism from the empty coalgebra to any T-coalgebra.

2.2 Paths and Bisimulations

If (A, α) and (B, β) are T-coalgebras, then a relation $R \subseteq A \times B$ is a *T-bisimulation* from α to β if there exists a transition structure $\rho : R \to TR$ on R such that the projections from R to A and B are coalgebraic morphisms from ρ to α and β, i.e. the following diagram commutes:



We may say that coalgebra β is the *image* of the bisimulation, or is the image of α under the bisimulation, if R is a surjective relation, i.e. every member of B is in the image of R, or equivalently π_2 is a surjective function. Dually, α is the *domain* of the bisimulation if R is a *total* relation, i.e. $\mathrm{Dom}\,R = A$, or equivalently π_1 is surjective.

A function $f : A \to B$ is a morphism from α to β iff its graph $\{(a, f(a)) : a \in A\}$ is a bisimulation from α to β [36, Theorem 2.5]: a morphism is essentially a functional bisimulation. When $\mathrm{Dom}\,\alpha \subseteq \mathrm{Dom}\,\beta$, α is a subcoalgebra of β iff the identity relation on $\mathrm{Dom}\,\alpha$ is a bisimulation from α to β.

The above categorical definition of bisimulation appeared in [1]. It has a characterisation [18, 19] in terms of "liftings" of relations $R \subseteq A \times B$ to relations $R^T \subseteq TA \times TB$. This in turn was transformed in [10] to another characterisation of bisimulations that uses the idea of "paths" between functors, an idea introduced in [24, Section 6].

A *path* is a finite list of symbols of the kinds π_j, ε_j, ev_d. Write $p.q$ for the concatenation of lists p and q. The notation $T \overset{p}{\rightsquigarrow} S$ means that p is a path from functor T to functor S, and is defined by the following conditions

- $T \overset{\langle\rangle}{\rightsquigarrow} T$, where $\langle\rangle$ is the empty path.

- $T_1 \times T_2 \overset{\pi_j.p}{\rightsquigarrow} S$ whenever $T_j \overset{p}{\rightsquigarrow} S$, for $j = 1, 2$.

- $T_1 + T_2 \overset{\varepsilon_j.p}{\rightsquigarrow} S$ whenever $T_j \overset{p}{\rightsquigarrow} S$, for $j = 1, 2$.

- $T^D \overset{ev_d.p}{\rightsquigarrow} S$ for all $d \in D$ whenever $T \overset{p}{\rightsquigarrow} S$.

It is evident that for any path $T \rightsquigarrow S$, S is one of the components of T. Paths can be composed by concatenating lists: if $T_1 \overset{p}{\rightsquigarrow} T_2$ and $T_2 \overset{q}{\rightsquigarrow} T_3$, then $T_1 \overset{p.q}{\rightsquigarrow} T_3$.

A path $T \overset{p}{\rightsquigarrow} S$ induces a partial function $p_A : TA \circ\!\!\longrightarrow SA$ for each set A, defined by induction on the length of p as follows.

- $\langle\rangle_A : TA \circ\!\!\longrightarrow TA$ is the identity function id_{TA}, so is totally defined.

- $(\pi_j.p)_A = p_A \circ \pi_j$, the composition of $T_1A \times T_2A \xrightarrow{\pi_j} T_jA \overset{p_A}{\circ\!\!\longrightarrow} SA$.
  Thus $x \in \mathrm{Dom}\,(\pi_j.p)_A$ iff $\pi_j(x) \in \mathrm{Dom}\,p_A$.

- $(\varepsilon_j.p)_A = p_A \circ \varepsilon_j$, the composition of $T_1A + T_2A \overset{\varepsilon_j}{\circ\!\!\longrightarrow} T_jA \overset{p_A}{\circ\!\!\longrightarrow} SA$.
  Thus $x \in \mathrm{Dom}\,(\varepsilon_j.p)_A$ iff $x \in \mathrm{Dom}\,\varepsilon_j$ and $\varepsilon_j(x) \in \mathrm{Dom}\,p_A$.

- $(ev_d.p)_A = p_A \circ ev_d$, the composition of $(TA)^D \xrightarrow{ev_d} TA \overset{p_A}{\circ\!\!\longrightarrow} SA$.
  Thus $f \in \mathrm{Dom}\,(ev_d.p)_A$ iff $f(d) \in \mathrm{Dom}\,p_A$.

Concatenation of paths corresponds to composition of functions, in the sense that $(p.q)_A = q_A \circ p_A$.

A path $T \rightsquigarrow S$ is a *state path* if $S = \mathrm{Id}$, an *observation path* if $S = \bar{D}$ for some set D, and a *basic* path if it is either. A straightforward induction on the formation of functors shows that if T is a polynomial functor, A a set, and $x \in TA$, then there exists a basic path $T \overset{p}{\rightsquigarrow} S$ with $x \in \mathrm{Dom}\,p_A$.

A T-bisimulation can be characterised as a relation that is "preserved" by the partial functions induced by state and observation paths from T. To explain this we adopt the convention that whenever we write "$f(x)\,Q\,g(y)$" for some relation Q and some partial functions f and g we mean that $f(x)$ is defined iff $g(y)$ is defined, and $(f(x), g(y)) \in Q$ when they are both defined.

Theorem 2.1 [10, Theorem 5.7]
If $A \xrightarrow{\alpha} TA$ and $B \xrightarrow{\beta} TB$ are coalgebras for a polynomial functor T, then a relation $R \subseteq A \times B$ is a T-bisimulation if, and only if, xRy implies

(1) *for all state paths $T \overset{p}{\rightsquigarrow} \mathrm{Id}$, $p_A(\alpha(x))\,R\,p_B(\beta(y))$; and*

(2) *for all observation paths $T \overset{p}{\rightsquigarrow} \bar{D}$, $p_A(\alpha(x)) = p_B(\beta(y))$.* □

The inverse of a bisimulation is a bisimulation, and the union of any collection of bisimulations from α to β is a bisimulation [36, Section 5]. Hence there is a largest bisimulation from α to β, which is called *bisimilarity.* We denote this by $\sim$. States x and y are *bisimilar*, $x \sim y$, when xRy for some bisimulation R between α and β. This is intended to capture the notion that x and y are observationally indistinguishable.

2.3 Types, Terms, and Formulas

Fix a set $\mathbb{O}$ of symbols called *observable types*, and a collection $\{[\![o]\!] : o \in \mathbb{O}\}$ of non-empty sets indexed by $\mathbb{O}$. $[\![o]\!]$ is the *denotation* of o, and its members are called *observable elements*, or *constants* of type o. The set of *types over* $\mathbb{O}$, or $\mathbb{O}$*-types*, is the smallest set $\mathbb{T}$ such that $\mathbb{O} \subseteq \mathbb{T}$, $\mathsf{St} \in \mathbb{T}$ and

(1) if $\sigma_1, \sigma_2 \in \mathbb{T}$ then $\sigma_1 \times \sigma_2,\ \sigma_1 + \sigma_2 \in \mathbb{T}$;

(2) if $\sigma \in \mathbb{T}$ and $o \in \mathbb{O}$, then $o \Rightarrow \sigma \in \mathbb{T}$.

A *subtype* of an $\mathbb{O}$-type τ is any type that occurs in the formation of τ. St is a type symbol that will denote the state set of a given coalgebra. The symbol "o" will always be reserved for members of $\mathbb{O}$. $o \Rightarrow \sigma$ is a power type: such types will always have an observable exponent. A type is *rigid* if it does not have St as a subtype. The set of rigid types is thus the smallest set that includes $\mathbb{O}$ and satisfies (1) and (2).

Each $\mathbb{O}$-type σ determines a polynomial functor $|\sigma| : \mathbf{Set} \to \mathbf{Set}$. For $o \in \mathbb{O}$, $|o|$ is the constant functor $\bar{D}$ where $D = [\![o]\!]$; $|\mathsf{St}|$ is the identity functor Id; and inductively

$$|\sigma_1 \times \sigma_2| = |\sigma_1| \times |\sigma_2|, \quad |\sigma_1 + \sigma_2| = |\sigma_1| + |\sigma_2|, \quad |o \Rightarrow \sigma| = |\sigma|^{[\![o]\!]}.$$

For denotations of types, we write $[\![\sigma]\!]_A$ for the set $|\sigma|A$. Thus we have $[\![o]\!]_A = [\![o]\!]$, $[\![\mathsf{St}]\!]_A = A$,

$$\begin{aligned} [\![\sigma_1 \times \sigma_2]\!]_A &= [\![\sigma_1]\!]_A \times [\![\sigma_2]\!]_A \\ [\![\sigma_1 + \sigma_2]\!]_A &= [\![\sigma_1]\!]_A + [\![\sigma_2]\!]_A \\ [\![o \Rightarrow \sigma]\!]_A &= [\![\sigma]\!]_A{}^{[\![o]\!]}. \end{aligned}$$

If σ is a rigid type then $|\sigma|$ is a constant functor, so $[\![\sigma]\!]_A$ is a fixed set whose definition does not depend on A and may be written $[\![\sigma]\!]$. A τ*-coalgebra* is a coalgebra (A, α) for the functor $|\tau|$, i.e. α is a function of the form $A \to [\![\tau]\!]_A$.

To define *terms* we fix a denumerable set *Var* of *variables* and define a *context* to be a finite (possible empty) list

$$\Gamma = (v_1 : \sigma_1, \ldots, v_n : \sigma_n)$$

of assignments of $\mathbb{O}$-types σ_i to variables v_i, with the proviso that $v_1, \ldots, v_n$ are all distinct. Γ is a *rigid* context if all of the σ_i's are rigid types. Concatenation of lists Γ and Γ' with disjoint sets of variables is written Γ, Γ'. A *term-in-context* is an expression of the form

$$\Gamma \rhd M : \sigma,$$

which signifies that M is a "raw" term of type σ in context Γ. This may be abbreviated to $\Gamma \rhd M$ if the type of the term is understood. If $\sigma \in \mathbb{O}$, then the term is *observable.*

In the semantics to follow in Section 2.4, a term $\Gamma \rhd M : \sigma$ is assigned a value at each state of a coalgebra, relative to an assigment to values to the variables appearing in Γ. The value of the term is an element of $[\![\sigma]\!]_A$.

Figure 2 gives axioms that legislate certain *base terms* into existence, and rules for generating new terms from given ones. Axiom (Con) states that an observable element is a constant term of its type, while the raw term s in axiom (St) is a special parameter which will be interpreted as the "current" state in a coalgebra. The rules for products, coproducts and powers are the standard ones for introduction and transformation of terms of those types. The raw term in the consequent of rule (Case) is sometimes abbreviated to $\mathsf{case}(N, M_1, M_2)$.

Bindings of variables in raw terms occur in lambda-abstractions and case terms: the v in the consequent of rule (Abs) and the v_j's in the consequent of (Case) are bound in those terms. It is readily shown that in any term $\Gamma \rhd M$, all free variables of M appear in the list Γ. A *ground* term is one of the form $\emptyset \rhd M : \sigma$, which may be abbreviated to $M : \sigma$, or just to the raw term M. Thus a ground term has no free variables. Note that a ground term may contain the state parameter s, which behaves nonetheless as a variable in that it

Axioms

(Var) $\dfrac{v \in Var}{v : \sigma \rhd v : \sigma}$ (Con) $\dfrac{c \in [\![o]\!]}{\emptyset \rhd c : o}$ (St) $\dfrac{}{\emptyset \rhd \mathsf{s} : \mathsf{St}}$

Weakening

(Weak) $\dfrac{\Gamma, \Gamma' \rhd M : \sigma}{\Gamma, v : \sigma', \Gamma' \rhd M : \sigma}$ where v does not occur in Γ or Γ'.

Product Types

(Proj$_1$) $\dfrac{\Gamma \rhd M : \sigma_1 \times \sigma_2}{\Gamma \rhd \pi_1 M : \sigma_1}$ (Proj$_2$) $\dfrac{\Gamma \rhd M : \sigma_1 \times \sigma_2}{\Gamma \rhd \pi_2 M : \sigma_2}$

(Pair) $\dfrac{\Gamma \rhd M_1 : \sigma_1 \qquad \Gamma \rhd M_2 : \sigma_2}{\Gamma \rhd \langle M_1, M_2 \rangle : \sigma_1 \times \sigma_2}$

Coproduct Types

(In$_1$) $\dfrac{\Gamma \rhd M : \sigma_1}{\Gamma \rhd \iota_1 M : \sigma_1 + \sigma_2}$ (In$_2$) $\dfrac{\Gamma \rhd M : \sigma_2}{\Gamma \rhd \iota_2 M : \sigma_1 + \sigma_2}$

(Case) $\dfrac{\Gamma \rhd N : \sigma_1 + \sigma_2 \qquad \Gamma, v_1 : \sigma_1 \rhd M_1 : \sigma \qquad \Gamma, v_2 : \sigma_2 \rhd M_2 : \sigma}{\Gamma \rhd \mathsf{case}\ N\ \mathsf{of}\ [\iota_1 v_1 \mapsto M_1 \,|\, \iota_2 v_2 \mapsto M_2] : \sigma}$

Power Types

(App) $\dfrac{\Gamma \rhd M : o \Rightarrow \sigma \qquad \Gamma \rhd N : o}{\Gamma \rhd M \cdot N : \sigma}$ (Abs) $\dfrac{\Gamma, v : o \rhd M : \sigma}{\Gamma \rhd (\lambda v.M) : o \Rightarrow \sigma}$

Figure 2: Axioms and Rules for Generating Terms

Equations

$$(\text{Eq})\quad \frac{\Gamma \rhd M_1 : \sigma \qquad \Gamma \rhd M_2 : \sigma}{\Gamma \rhd M_1 \approx M_2}$$

Weakening

$$(\text{Weak})\quad \frac{\Gamma, \Gamma' \rhd \varphi}{\Gamma, v : \sigma', \Gamma' \rhd \varphi} \qquad \text{where } v \text{ does not occur in } \Gamma \text{ or } \Gamma'.$$

Connectives

$$(\text{Neg})\quad \frac{\Gamma \rhd \varphi}{\Gamma \rhd \neg\varphi} \qquad\qquad (\text{Con})\quad \frac{\Gamma \rhd \varphi_1 \qquad \Gamma \rhd \varphi_2}{\Gamma \rhd \varphi_1 \wedge \varphi_2}$$

Figure 3: Formation Rules for Formulas

can denote any member of $\mathrm{Dom}\,\alpha$, as will be seen in the semantics presented in Section 2.4. There exist ground terms of every type, as may be seen by induction on type formation.

A term is defined to be *rigid* if its context is rigid. This entails that any free variable of the term is assigned a rigid type by Γ, so its type is formed without use of St. Of course all ground terms are rigid.

τ-Terms

For a given $\mathbb{O}$-type τ, a *τ-term* is any term that can be generated by the axioms and rules of Figure 2 together with the additional rule

$$(\tau\text{-Tr})\quad \frac{\Gamma \rhd M : \mathsf{St}}{\Gamma \rhd \mathsf{tr}(M) : \tau}\ .$$

Note that from this rule and the axiom (St) we can derive the ground τ-term

$$\emptyset \rhd \mathsf{tr}(\mathsf{s}) : \tau.$$

The symbol tr will denote the transition structure of a τ-coalgebra $A \xrightarrow{\alpha} [\![\,\tau\,]\!]_A$. If M is a ground term of type St, then the value of the term $\mathsf{tr}(M)$ at a state x is obtained by applying α to the value of M at x. Since the parameter s is to be interpreted as the "current state", its value at state x is just x, so the value of $\mathsf{tr}(\mathsf{s})$ at x is $\alpha(x)$. Hence the term $\mathsf{tr}(\mathsf{s})$ denotes the transition function α itself.

τ-Formulas

An *equation-in-context* has the form $\Gamma \rhd M_1 \approx M_2$ where $\Gamma \rhd M_1$ and $\Gamma \rhd M_2$ are terms *of the same type*. A *formula-in-context* has the form $\Gamma \rhd \varphi$, with the expression φ being constructed from equations $M_1 \approx M_2$ by propositional connectives. Formation rules for formulas are given in Figure 3, using the connectives $\neg$ and $\wedge$. The other standard connectives $\vee$, $\rightarrow$, and $\leftrightarrow$ can be introduced as definitional abbreviations in the usual way. We may also write $(M_1 \not\approx M_2)$ for $\neg(M_1 \approx M_2)$, and will use the symbol $\perp$ to stand for any formula of the form $\Gamma \rhd \varphi \wedge \neg\varphi$.

A formula $\emptyset \rhd \varphi$ with empty context is *ground*, and may be abbreviated to φ. A *rigid* formula is one whose context is rigid. A *τ-formula* is one that is generated by using only τ-terms as premisses in the rule (Eq). An *observable* formula is one that uses only terms of observable type in forming its component equations.

2.4 Semantics of Terms and Formulas

A τ-coalgebra $\alpha : A \to |\tau| A$ interprets types σ and contexts $\Gamma = (v_1 : \sigma_1, \ldots, v_n : \sigma_n)$ by putting $[\![\,\sigma\,]\!]_\alpha = |\sigma|(\mathrm{Dom}\,\alpha) = [\![\,\sigma\,]\!]_A$, and

$$[\![\,\Gamma\,]\!]_\alpha = [\![\,\sigma_1\,]\!]_\alpha \times \cdots \times [\![\,\sigma_n\,]\!]_\alpha$$

(so $[\![\emptyset]\!]_\alpha$ is the empty product 1). The *denotation* of each τ-term $\Gamma \rhd M : \sigma$, relative to the coalgebra α, is a function

$$[\![\Gamma \rhd M : \sigma]\!]_\alpha : A \times [\![\Gamma]\!]_\alpha \longrightarrow [\![\sigma]\!]_\alpha,$$

defined by induction on the formation of terms. For empty contexts,

$$A \times [\![\emptyset]\!]_\alpha = A \times 1 \cong A,$$

so we replace $A \times [\![\emptyset]\!]_\alpha$ by A itself and interpret a ground term $\emptyset \rhd M : \sigma$ as a function $A \to [\![\sigma]\!]_\alpha$. The definition of denotations is as follows.

Var:

$[\![v : \sigma \rhd v : \sigma]\!]_\alpha : A \times [\![\sigma]\!]_\alpha \to [\![\sigma]\!]_\alpha$ is the right projection function.

Con:

$[\![\emptyset \rhd c : o]\!]_\alpha : A \to [\![o]\!]$ is the constant function with value c.

St:

$[\![\emptyset \rhd \mathsf{s} : \mathsf{St}]\!]_\alpha : A \to [\![\mathsf{St}]\!]_\alpha$ is the identity function $A \to A$.

τ-Tr:

$[\![\Gamma \rhd \mathsf{tr}(M) : \tau]\!]_\alpha : A \times [\![\Gamma]\!]_\alpha \to [\![\tau]\!]_\alpha$ is the composition of the functions

$$A \times [\![\Gamma]\!]_\alpha \xrightarrow{[\![\Gamma \rhd M : \mathsf{St}]\!]_\alpha} A \xrightarrow{\alpha} [\![\tau]\!]_\alpha.$$

Weak:

$[\![\Gamma, v : \sigma', \Gamma' \rhd M : \sigma]\!]_\alpha$ is the composition of $[\![\Gamma, \Gamma' \rhd M : \sigma]\!]_\alpha$ with the projection

$$A \times [\![\Gamma]\!]_\alpha \times [\![\sigma']\!]_\alpha \times [\![\Gamma']\!]_\alpha \longrightarrow A \times [\![\Gamma]\!]_\alpha \times [\![\Gamma']\!]_\alpha.$$

Proj$_j$:

$[\![\Gamma \rhd \pi_j M : \sigma_j]\!]_\alpha$ is the composition of

$$A \times [\![\Gamma]\!]_\alpha \xrightarrow{[\![\Gamma \rhd M : \sigma_1 \times \sigma_2]\!]_\alpha} [\![\sigma_1]\!]_\alpha \times [\![\sigma_2]\!]_\alpha \xrightarrow{\pi_j} [\![\sigma_j]\!]_\alpha.$$

Pair:

$[\![\Gamma \rhd \langle M_1, M_2\rangle : \sigma_1 \times \sigma_2]\!]_\alpha$ is the pairing function

$$A \times [\![\Gamma]\!]_\alpha \xrightarrow{\langle [\![\Gamma \rhd M_1 : \sigma_1]\!]_\alpha, [\![\Gamma \rhd M_2 : \sigma_2]\!]_\alpha \rangle} [\![\sigma_1]\!]_\alpha \times [\![\sigma_2]\!]_\alpha.$$

Inj$_j$:

$[\![\Gamma \rhd \iota_j M : \sigma_1 + \sigma_2]\!]_\alpha$ is the composition of

$$A \times [\![\Gamma]\!]_\alpha \xrightarrow{[\![\Gamma \rhd M : \sigma_j]\!]_\alpha} [\![\sigma_j]\!]_\alpha \xrightarrow{\iota_j} [\![\sigma_1]\!]_\alpha + [\![\sigma_2]\!]_\alpha.$$

Case:

This is easier to describe at the function-value level. For $x \in A$ and $\gamma \in [\![\Gamma]\!]_\alpha$, let

$$[\![\Gamma \rhd N : \sigma_1 + \sigma_2]\!]_\alpha(x, \gamma) = \iota_j(a) \in [\![\sigma_1]\!]_\alpha + [\![\sigma_2]\!]_\alpha,$$

which holds for a unique j and $a \in [\![\sigma_j]\!]_\alpha$. Then the element

$$[\![\Gamma \rhd \mathsf{case}\ N\ \mathsf{of}\ [\iota_1 v_1 \mapsto M_1 \mid \iota_2 v_2 \mapsto M_2] : \sigma]\!]_\alpha(x, \gamma)$$

of $[\![\sigma]\!]_\alpha$ is defined to be

$$[\![\Gamma, v_j : \sigma_j \rhd M_j : \sigma]\!]_\alpha(x, \gamma, a).$$

In other words: if $[\![\Gamma \rhd N]\!]_\alpha(x, \gamma) \in \mathrm{Dom}\, \varepsilon_j$, then

$$[\![\Gamma \rhd \mathsf{case}(N, M_1, M_2)]\!]_\alpha(x, \gamma) = [\![\Gamma, v_j : \sigma_j \rhd M_j]\!]_\alpha(x, \gamma, \varepsilon_j [\![\Gamma \rhd N]\!]_\alpha(x, \gamma)).$$

App:

$[\![\Gamma \rhd M \cdot N : \sigma]\!]_\alpha(x, \gamma)$ is the element of $[\![\sigma]\!]_\alpha$ obtained by evaluating the function

$$[\![\Gamma \rhd M : o \Rightarrow \sigma]\!]_\alpha(x, \gamma) : [\![o]\!] \longrightarrow [\![\sigma]\!]_\alpha$$

at $[\![\Gamma \rhd N : o]\!]_\alpha(x, \gamma) \in [\![o]\!]$.

Abs:

$[\![\Gamma \rhd (\lambda v.M) : o \Rightarrow \sigma]\!]_\alpha(x, \gamma)$ is the function $[\![o]\!] \to [\![\sigma]\!]_\alpha$ given by

$$a \mapsto [\![\Gamma, v : o \rhd M : \sigma]\!]_\alpha(x, \gamma, a).$$

This completes the inductive definition of $[\![\Gamma \rhd M : \sigma]\!]_\alpha$.

Substitution of Terms

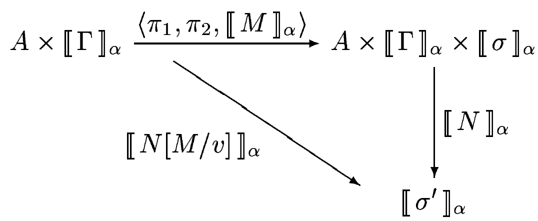
In working with this system it becomes essential to have available the operation $N[M/v]$ of substituting the raw term M for free occurrences of the variable v in N. The following rule is derivable:

$$\text{(Subst)} \quad \frac{\Gamma \rhd M : \sigma \qquad \Gamma, v : \sigma \rhd N : \sigma'}{\Gamma \rhd N[M/v] : \sigma'}$$

The semantics of terms obeys the basic principle that substitution is interpreted as *composition* of denotations [31, 2.2]. Because of the special role of the state set A, this takes the form

$$[\![\Gamma \rhd N[M/v]]\!]_\alpha = [\![\Gamma, v : \sigma \rhd N]\!]_\alpha \circ \langle \pi_1, \pi_2, [\![\Gamma \rhd M]\!]_\alpha \rangle,$$

so that the following diagram commutes:

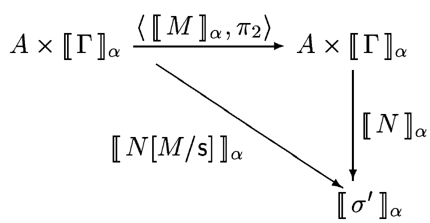


Substitution for the State Parameter

It is also possible to make substitutions $N[M/\mathsf{s}]$ for the state parameter s according to the derivable rule

$$\text{(s-Subst)} \quad \frac{\Gamma \rhd M : \mathsf{St} \qquad \Gamma \rhd N : \sigma'}{\Gamma \rhd N[M/\mathsf{s}] : \sigma'}$$

with the semantics $[\![\Gamma \rhd N[M/\mathsf{s}]]\!]_\alpha = [\![\Gamma \rhd N]\!]_\alpha \circ \langle [\![\Gamma \rhd M]\!]_\alpha, \pi_2 \rangle$:



For ground terms ($\Gamma = \emptyset$), this takes the simple form

$$[\![N[M/\mathsf{s}]]\!]_\alpha = [\![N]\!]_\alpha \circ [\![M]\!]_\alpha.$$

Semantics of Formulas

A τ-equation $\Gamma \rhd M_1 \approx M_2$ is said to be *valid* in coalgebra α if the α-denotations $[\![\Gamma \rhd M_1]\!]_\alpha$ and $[\![\Gamma \rhd M_2]\!]_\alpha$ of the terms $\Gamma \rhd M_j$ are identical. More generally we introduce a satisfaction relation

$$\alpha, x, \gamma \models \Gamma \rhd \varphi,$$

for τ-formulas in τ-coalgebras, which expresses that $\Gamma \rhd \varphi$ is *satisfied*, or *true*, in α at state x under the value-assigment $\gamma \in [\![\Gamma]\!]_\alpha$ to the variables of context Γ. This is defined inductively by

$$\begin{array}{lll} \alpha, x, \gamma \models \Gamma \rhd M_1 \approx M_2 & \text{iff} & [\![\Gamma \rhd M_1]\!]_\alpha(x,\gamma) = [\![\Gamma \rhd M_2]\!]_\alpha(x,\gamma), \\ \alpha, x, \gamma \models \Gamma \rhd \neg\varphi & \text{iff} & \text{not } \alpha, x, \gamma \models \Gamma \rhd \varphi, \\ \alpha, x, \gamma \models \Gamma \rhd \varphi_1 \wedge \varphi_2 & \text{iff} & \alpha, x, \gamma \models \Gamma \rhd \varphi_1 \text{ and } \alpha, x, \gamma \models \Gamma \rhd \varphi_2. \end{array}$$

$\Gamma \rhd \varphi$ is *true at x*, written $\alpha, x \models \Gamma \rhd \varphi$, if $\alpha, x, \gamma \models \Gamma \rhd \varphi$ for all $\gamma \in [\![\Gamma]\!]_\alpha$. α is a *model of* $\Gamma \rhd \varphi$, written $\alpha \models \Gamma \rhd \varphi$, if $\alpha, x, \models \Gamma \rhd \varphi$ for all states $x \in \mathrm{Dom}\,\alpha$. In that case we also say that $\Gamma \rhd \varphi$ is *valid in* the coalgebra α.

These definitions imply that if α is the *empty* τ-coalgebra, then α is a model of every τ-formula, since there is no state in α at which a formula can be false. In particular, $\alpha \models \perp$ when α is empty.

The following result is proven in [10, Section 5].

Theorem 2.2 *The class $\{\alpha : \alpha \models \Gamma \rhd \varphi\}$ of all models of an observable formula is closed under domains and images of bisimulations, including domains and images of morphisms as well as subcoalgebras. If $\Gamma \rhd \varphi$ is rigid and observable, then its class of models is also closed under disjoint unions.* □

The notation $Mod\,\Phi$ will used for the class of all models of a set Φ of formulas, and $Mod\,\varphi$ in the case that Φ consists of a single formula φ.

An Example: Streams of Characters

There are many examples of coalgebraic presentations of data structures to be found in the literature, in such sources as [32, 23, 22, 36, 26, 25]. We now develop an example of this kind, motivated by ideas from [23, 26], to illustrate features of the syntax and semantics just defined.

Imagine a simple game machine with a display screen and two buttons labeled play and next. The game starts with a blank screen. Pushing play causes a character from some character set C to be printed on the screen. Then pushing next causes the machine to move to a new state, from which another play action can be performed. Repeating these actions results in a string of characters being printed on the screen. But there are some states in which pushing play causes a "game over" message to be printed. In that case, pushing next causes the screen to be cleared for another game.

If A is the set of possible states of the machine, then its behaviour can be represented by functions

$$play : A \to C + \{over\}, \qquad next : A \to A, \qquad blank : A \to \{true, false\}.$$

play assigns to each state the character or *over* message resulting from pushing play. *next* gives the new state produced by pushing next, and *blank* assigns a truth value to a state according to whether or not the screen is clear in that state. These functions combine into a coalgebra

$$A \xrightarrow{\alpha} (C + \{over\}) \times (A \times \{true, false\}),$$

with $play = \pi_1 \circ \alpha$, $next = \pi_1 \circ (\pi_2 \circ \alpha)$, and $blank = \pi_2 \circ (\pi_2 \circ \alpha)$.

Let $\mathbb{O} = \{\mathsf{data}, 1, \mathsf{bool}\}$, with $[\![\mathsf{data}]\!] = C$, $[\![1]\!] = \{over\}$ and $[\![\mathsf{bool}]\!] = \{true, false\}$. Define the type τ to be $(\mathsf{data} + 1) \times (\mathsf{St} \times \mathsf{bool})$. Then (A, α) above is a τ-coalgebra.

Now for any ground τ-term M of type St, let $\mathsf{play}(M)$ be the term $\pi_1\mathsf{tr}(M)$ of type $\mathsf{data} + 1$. The denotation $[\![\mathsf{play}(M)]\!]_\alpha$ is the function $x \mapsto play([\![M]\!]_\alpha(x))$. In particular, $[\![\mathsf{play(s)}]\!]_\alpha$ is just the function *play*. Similarly we define $\mathsf{next}(M)$ to be the term $\pi_1\pi_2\mathsf{tr}(M)$

of type St, and $\mathsf{blank}(M)$ to be the term $\pi_2\pi_2\mathsf{tr}(M)$ of type bool, so that $\mathsf{next}(\mathsf{s})$ denotes the function *next* and $\mathsf{blank}(\mathsf{s})$ denotes *blank*.

Since $\mathsf{next}(M)$ also has type St, we can iterate its formation to obtain terms $\mathsf{next}^n(M)$ denoting the iterations of the function *next*, by putting $\mathsf{next}^1(M) = \mathsf{next}(M)$ and inductively $\mathsf{next}^{n+1}(M) = \mathsf{next}(\mathsf{next}^n(M))$.

We will write M-blank for the ground observable equation $(\mathsf{blank}(M) \approx \mathit{true})$, which is satisfied at state x iff $\mathit{blank}([\![M]\!]_\alpha(x)) = \mathit{true}$.

Notice that $\mathsf{play}(M)$ is not an observable term, since its type is $\mathsf{data}+1$, rather than one of the observable types data and 1. But we can describe which component of the disjoint union $C + \{\mathit{over}\}$ a value $\mathit{play}(x)$ belongs to, by making use of case terms. Let $\mathsf{test}(M)$ be the term

$$\mathsf{case}\ \mathsf{play}(M)\ \mathsf{of}\ [\iota_1 v_1 \mapsto \mathit{true} \mid \iota_2 v_2 \mapsto \mathit{false}],$$

which is of type bool. Then $[\![\mathsf{test}(M)]\!]_\alpha(x)$ is equal to *true* if $\mathit{play}([\![M]\!]_\alpha(x)) = \iota_1 c$ for some $c \in C$, and equal to *false* if $\mathit{play}([\![M]\!]_\alpha(x)) = \iota_2 \mathit{over}$. Thus we define M-live to be the formula $(\mathsf{test}(M) \approx \mathit{true})$ and M-over to be $(\mathsf{test}(M) \approx \mathit{false})$. These are ground observable equations. M-over asserts that the game will be over at the state denoted by M, while M-live asserts the opposite.

These formulas can be used to express constraints we might want a τ-coalgebra to satisfy in order to accurately model our game machine, such as

$$M\text{-}\mathsf{over} \to \mathsf{next}(M)\text{-}\mathsf{blank}$$
$$M\text{-}\mathsf{live} \to \neg\, \mathsf{next}(M)\text{-}\mathsf{blank}$$
$$M\text{-}\mathsf{blank} \to M\text{-}\mathsf{live}.$$

Operating the machine produces a stream of characters which will terminate if a state is reached that satisfies the formula s-over. If such a state is not reached, then an infinite stream will be generated. To express that this happens, let φ_n be the ground observable formula

$$\mathsf{s}\text{-}\mathsf{blank} \to \mathsf{next}^n(\mathsf{s})\text{-}\mathsf{live},$$

which asserts that the game does not end after n plays. Let $\Phi_\omega = \{\varphi_n : n \geq 1\}$. Then α is a model of Φ_ω iff all plays of the game are endless.

It is relatively straightforward to show that the class $\mathit{Mod}\,\Phi_\omega$ of all models of the infinite set Φ_ω is not equal to $\mathit{Mod}\,\Phi$ where Φ is any finite subset of Φ_ω. Fix an element $d \in C$, and for each $n \geq 1$ define a τ-coalgebra (A_n, α_n) that has $A_n = \{0, \ldots, n\}$, with α_n given by the functions next_n, blank_n, play_n defined as follows:

(1) $\mathit{next}_n(x) = x + 1$ for $x < n$, with $\mathit{next}_n(n) = 0$;

(2) $\mathit{blank}_n(x) = \mathit{true}$ iff $x = 0$.

(3) $\mathit{play}_n(x) = \iota_1 d$ if $x < n$, while $\mathit{play}_n(n) = \iota_2\, \mathit{over}$.

s-blank is true only at 0, while s-live is true at all states except n. So φ_n is false at 0 in α_n, but φ_k is valid in α_n for all $k < n$. Thus given any finite number of positive integers $n_1, \ldots, n_k$, choosing any n larger than all of them gives

$$\alpha_n \in \mathit{Mod}\,\{\varphi_{n_1}, \ldots, \varphi_{n_k}\} - \mathit{Mod}\,\Phi_\omega.$$

At the end of Section 7 we will apply an ultraproduct construction to these α_n's to show that there is no τ-formula φ whatsoever with $\mathit{Mod}\,\varphi = \mathit{Mod}\,\Phi_\omega$.

The kind of use made of case terms in this example plays a vital role in [10] in showing that path functions p_A and extraction functions ε_j are term-definable, and ultimately in showing that bisimilarity of states is characterized by their assigning the same values to observable terms (see Theorem 3.2 below). As the example demonstrates, these results depend on the presence of at least one observable type o (like bool) for which $[\![o]\!]$ has at least two elements. The next section explains this further.

3 Defining Path Action and Bisimilarity

The action of a path function is definable by a (ground) term, in the following sense.

Lemma 3.1 (Path Lemma) [10, Theorem 6.1]
For any path $|\tau| \overset{p}{\rightsquigarrow} |\sigma|$ *and variable* v *there exists a* tr*-free* τ*-term of the form*

$$v : \tau \rhd \bar{p} : \sigma$$

such that for any τ*-coalgebra* (A, α) *and any* $x \in A$*, if* $\alpha(x) \in \operatorname{Dom} p_A$ *then*

$$p_A(\alpha(x)) = [\![\bar{p}[\mathsf{tr}(\mathsf{s})/v]]\!]_\alpha(x). \qquad \square$$

Note that by the substitution rule (Subst), $\bar{p}[\mathsf{tr}(\mathsf{s})/v]$ is a ground term of type σ, since $\mathsf{tr}(\mathsf{s})$ is a ground term of type τ . The term function $[\![\bar{p}[\mathsf{tr}(\mathsf{s})/v]]\!]_\alpha$ has domain A, and so may not be identical to $p_A \circ \alpha$ if p_A is partial. This is only an issue when the path p includes an extraction symbol ε_j (for otherwise p_A is total), but use of case allows the construction of observable terms that "discriminate" between the two summands of a coproduct $[\![\tau_1]\!]_A + [\![\tau_2]\!]_A$ and determine whether $p_A(\alpha(x))$ is defined [10, Section 6]. For this to work it is necessary to assume that there is available at least one observable type o that is *non-trivial* in the sense that $[\![o]\!]$ has at least two distinct members. This is a plausible assumption in dealing with notions that are to be discriminated by observable behaviour. Define a relation $\equiv_{\alpha\beta}$ between the state sets of two τ-coalgebras by putting

$$x \equiv_{\alpha\beta} y \text{ iff every } \textit{ground observable} \text{ term } M \text{ has } [\![M]\!]_\alpha(x) = [\![M]\!]_\beta(y).$$

If τ has at least one non-trivial observable subtype, and $x \equiv_{\alpha\beta} y$, then for any path $|\tau| \overset{p}{\rightsquigarrow} |\sigma|$, $\alpha(x) \in \operatorname{Dom} p_A$ iff $\beta(y) \in \operatorname{Dom} p_B$ [10, Corollary 6.3]. This observation, together with Theorem 2.1 and further use of the Path Lemma leads to a proof that $\equiv_{\alpha\beta}$ is a bisimulation from α to β [10, Lemma 7.1]. Moreover it proves to be the largest bisimulation, giving a logical definition of bisimilarity. The precise situation is as follows.

Theorem 3.2 [10, Theorem 7.2]
Let (A, α) *and* (B, β) *be* τ*-coalgebras, where* τ *has at least one non-trivial observable subtype. Then for any* $x \in A$ *and* $y \in B$*, the following are equivalent:*

(1) x *and* y *are bisimilar:* $x \sim y$.

(2) $\alpha, x \models \Gamma \rhd \varphi$ *iff* $\beta, y \models \Gamma \rhd \varphi$ *for all rigid observable formulas* $\Gamma \rhd \varphi$.

(3) $\alpha, x \models M \approx N$ *implies* $\beta, y \models M \approx N$ *for all ground observable terms* M *and* N.

(4) $[\![M]\!]_\alpha(x) = [\![M]\!]_\beta(y)$ *for all ground observable terms* M*, i.e.* $x \equiv_{\alpha\beta} y$.

4 Ultraproducts and Ultrapowers

The theory of ultraproducts and ultrapowers plays a fundamental role in classical model theory [6, 9]. We are going to make a typical use of ultrapowers, and to motivate it we first review the nature of these classical constructions.

Let $\{\mathcal{A}_i : i \in I\}$ be a set of models for some first-order language $\mathcal{L}$, and U an ultrafilter on I. The ultraproduct $\Pi_U \mathcal{A}_i$ of the $\mathcal{A}_i$'s over U is a quotient of the ordinary direct product $\Pi_I \mathcal{A}_i$ that is obtained by identifying any two functions $f, g \in \Pi_I \mathcal{A}_i$ whose *agreement set*

$$\{i \in I : f(i) = g(i)\}$$

is a member of U. If f^U is the set of g's that are identified with f in this way, then $\Pi_U \mathcal{A}_i$ is based on the set $\{f^U : f \in \Pi_I \mathcal{A}_i\}$.

The ultrafilter U may be informally viewed as a collection of "large" subsets of I: each member of U contains "almost all" members of I. Thus $\Pi_U \mathcal{A}_i$ identifies any two members of $\Pi_I \mathcal{A}_i$ that agree "almost everywhere".

The importance of this construction derives from the fact that the first-order definable properties possessed by $\Pi_U \mathcal{A}_i$ are just those that are possessed by almost all of the $\mathcal{A}_i$'s: $\Pi_U \mathcal{A}_i$ is a model of any given $\mathcal{L}$-sentence φ iff almost all of the $\mathcal{A}_i$'s are models of φ, in the sense that the set

$$\{i \in I : \mathcal{A}_i \text{ is a model of } \varphi\}$$

belongs to U. This fact is a consequence of a stronger result, known as *Łoś's Theorem* [6, Section 5.2], about satisfaction of $\mathcal{L}$-formulas. If such a formula φ has free variables amongst $v_1, \ldots, v_n$, and $f_1^U, \ldots, f_n^U$ is a sequence of elements of the ultraproduct that interpret these variables, then φ is satisfied in $\Pi_U \mathcal{A}_i$ by this interpretation iff it is satisfied in almost all of the $\mathcal{A}_i$'s by the interpretation $f_1(i), \ldots, f_n(i)$. Symbolically, Łoś's Theorem can be written as the statement that

$$\Pi_U \mathcal{A}_i, f_1^U, \ldots, f_n^U \models \varphi \quad \text{if, and only if,} \quad \{i \in I : \mathcal{A}_i, f_1(i), \ldots, f_n(i) \models \varphi\} \in U.$$

Informally we can think of $\Pi_U \mathcal{A}_i$ as the "average" of all the structures $\mathcal{A}_i$, just as the mythical "average person" is someone whose characteristics are those possessed by most people.

In the case that the $\mathcal{A}_i$'s are all equal to a single model $\mathcal{A}$, the ultraproduct is called the ultra*power* of $\mathcal{A}$ over U. It is sometimes written as $\Pi_U \mathcal{A}$ or $\mathcal{A}/U$, but we find it convenient to use the notation $\mathcal{A}^U$ for it. Each element a of $\mathcal{A}$ determines the element $\bar{a}^U$ of $\mathcal{A}^U$, where $\bar{a} \in A^I$ is the constant function on I with value a. Then Łoś's Theorem implies that

$$\mathcal{A}^U, \bar{a}_1^U, \ldots, \bar{a}_n^U \models \varphi \quad \text{if, and only if,} \quad \mathcal{A}, a_1, \ldots, a_n \models \varphi.$$

If a and $\bar{a}^U$ are identified, $\mathcal{A}^U$ becomes an extension of the model $\mathcal{A}$, and indeed an *elementary* extension according to the last observation.

One important role played by ultrapowers, and the one to be played here, is to provide *saturated* extensions of given models. Here "saturation" refers to the idea that a model is "full of elements". There are various forms of this notion, and typically they assert that a model must contain an element meeting a certain description whenever that description is logically consistent with the description of the model itself. An ultrapower that has the particular form of saturation we need will be referred to as *enlarging* (see Section 6). This is based on the notion that a subset C of $\mathcal{A}$ can be enlarged to a subset of $\mathcal{A}^U$ by adding to it any member f^U of $\mathcal{A}^U$ for which almost all f-values are members of C, i.e. for which $\{i \in I : f(i) \in C\} \in U$. Such an f^U can be thought of as a "nonstandard member" of C. When C is defined in $\mathcal{A}$ by some property of elements that is specifiable by a first-order formula, then by Łoś's Theorem f^U will be a nonstandard member of C precisely when it itself has this same property in $\mathcal{A}^U$.

Now a collection $\mathcal{S}$ of subsets of A may be viewed as a collection of properties that is "logically consistent with the description of $\mathcal{A}$" when it has the *finite intersection property*, i.e. when any finite collection of sets from $\mathcal{S}$ has non-empty intersection. That means that any finite number of these properties are simultaneously satisfied by some element of $\mathcal{A}$. The ultrapower $\mathcal{A}^U$ is *enlarging* if the following holds:

> for any collection $\mathcal{S}$ of subsets of A with the finite intersection property there is an element of $\mathcal{A}^U$ that is a nonstandard member of every set in $\mathcal{S}$.

Turning now to the case of coalgebras, we immediately strike an obstacle which can be illustrated with the case of an ultrapower. If (A, α) is a τ-coalgebra, the classical ultrapower construction lifts α to a function α^U on the ultrapower A^U of the state set A. But this α^U is not in general a τ-coalgebra, as it has the form $A^U \to (|\tau|A)^U$ rather than $A^U \to |\tau|(A^U)$ (the reason why is explained below). To resolve this problem, certain points will be removed from A^U to reduce it to a subset A^+ on which a suitable τ-transition $\alpha^+ : A^+ \to |\tau|A^+$ can be constructed. Members of A^+ are are defined by a property involving denotations of observable terms, and α^+ is called the *observational ultrapower* of α. The construction of α^+ is the subject of this section, and in Section 5 we establish a version of Łoś's Theorem for α^+ that allows us to conclude that α and α^+ validate the same observable formulas.

Now when α^+ is enlarging, it is sufficiently full of elements (saturated) that the following holds (Theorem 6.1):

if β is a τ-coalgebra validating every ground observable formula that is valid in α, then for each state y of β there exists a state of α^+ that is bisimilar to y, so the bisimilarity relation from α^+ to β is surjective.

Indeed, to show that this holds it is enough, by Theorem 3.2, to show that α^+ has a state satisfying every ground observable equation that is satisfied by y. But by the enlarging property, this reduces to showing that the set of all such equations is logically consistent with α, in the sense that each finite set of such equations can be satisfied at some state of α.

The upshot is that a *logically* specified relationship between α and β implies the *structural* relationship that β is the image of α^+ under bisimilarity. This result leads to our structural characterization of classes of coalgebras definable by observable formulas (Theorem 7.1). It also allows us to show (in Theorem 6.2), that if α and β are logically indistinguishable by any ground observable formula then they have ultrapowers α^+ and β^+ that are observationally indistinguishable in the sense that each state of one is bisimilar to a state of the other, so any behaviour of either is represented in the other.

With this motivating account of our programme in mind, we turn to the formal development. Let U be an ultrafilter on a set I. Thus $\emptyset \notin U$ and U is closed under supersets and finite intersections and contains exactly one of J and $I-J$ for each $J \subseteq I$. U is *non-principal* if its intersection is empty. A non-principal ultrafilter does not contain any finite subset of I.

Extensive use will be made of "tuple" notation. A collection of sets A_i indexed by the members of I will be presented as the *I-tuple* $\langle A_i : i \in I\rangle$. A function f with domain I may be written as the I-tuple $\langle f(i) : i \in I\rangle$, or even more briefly as $\langle f(i)\rangle$. The latter notation is particularly convenient when f is defined by some complex expression for $f(i)$.

There is an equivalence relation $=_U$ on the direct product $\Pi_I A_i$, defined by

$$f =_U g \text{ iff } \{i \in I : f(i) = g(i)\} \in U.$$

Each $f \in \Pi_I A_i$ has the equivalence class $f^U = \{g \in \Pi_I A_i : f =_U g\}$. The quotient set

$$\Pi_U A_i = \{f^U : f \in \Pi_I A_i\}$$

is called the *ultraproduct of* $\langle A_i : i \in I\rangle$ *over* U. f^U may be written as $\langle f(i) : i \in I\rangle^U$, or even as $\langle f(i)\rangle^U$.

In the case that the A_i's are all equal to a single set A, then the ultraproduct $\Pi_U A_i$ is called the *ultrapower of A over U*, and is written as A^U. A notation that will be useful for ultrapowers is to write $f \in_U C$, for $C \subseteq A$, when $\{i \in I : f(i) \in C\} \in U$. We may also safely write $f^U \in_U C$ in this case, since in general $f \in_U C$ iff $g \in_U C$ whenever $f =_U g$.

There is a natural injection $e_A : A \rightarrowtail A^U$ given by $e_A(a) = \bar{a}^U$, where the function $\bar{a} \in A^I$ has $\bar{a}(i) = a$ for all $i \in I$. The distinction between a and $\bar{a}^U$ is sometimes elided, allowing A to be identified with the subset $e_A(A)$ of A^U.

An I-tuple $\langle \theta_i : i \in I\rangle$ of n-ary functions of the form

$$\theta_i : A_{1i} \times \cdots \times A_{ni} \to B_i$$

has a *U-lifting* to the function

$$\theta^U : \Pi_U A_{1i} \times \cdots \times \Pi_U A_{ni} \to \Pi_U B_i,$$

given by

$$\theta^U(f_1^U, \ldots, f_n^U) = \langle \theta_i(f_1(i), \ldots, f_n(i)) : i \in I\rangle^U.$$

In the case $n = 1$, a family of maps $\theta_i : A_i \to B_i$ lifts to $\theta^U : \Pi_U A_i \to \Pi_U B_i$ where $\theta^U(f^U) = \langle \theta_i(f(i)) : i \in I\rangle^U$. This works also for *partial* $\theta_i : A_i \circ\!\!\longrightarrow B_i$, providing a partial U-lifting $\theta^U : \Pi_U A_i \circ\!\!\longrightarrow \Pi_U B_i$ in the same way, with the proviso that $f^U \in \operatorname{Dom}\theta^U$ precisely when $\{i \in I : f(i) \in \operatorname{Dom}\theta_i\} \in U$. Moreover, U-lifting commutes with functional composition: given also $\eta_i : B_i \circ\!\!\longrightarrow C_i$ we have $\langle \eta_i \circ \theta_i\rangle^U = \langle \eta_i\rangle^U \circ \langle \theta_i\rangle^U : \Pi_U A_i \circ\!\!\longrightarrow \Pi_U C_i$.

In the ultrapower case, $\theta : A \circ\!\!\longrightarrow B$ lifts to $\theta^U : A^U \circ\!\!\longrightarrow B^U$ where $\theta^U(f^U) = (\theta \circ f)^U$, with $f^U \in \mathrm{Dom}\,\theta^U$ precisely when $f \in_U \mathrm{Dom}\,\theta$.

Projections are preserved by U-liftings, but this can be interpreted in two ways. Given two collections $\langle A_{1i} : i \in I\rangle$ and $\langle A_{2i} : i \in I\rangle$, with associated projections $\pi_{ji} : A_{1i} \times A_{2i} \to A_{ji}$ (for $j = 1, 2$), let $\pi_j = \langle \pi_{ji} : i \in I\rangle$. Then π_j^U proves to be the projection $\Pi_U A_{1i} \times \Pi_U A_{2i} \to \Pi_U A_{ji}$, i.e. $\pi_j^U(x_1, x_2) = x_j$. But there is an equally important function of the form $\Pi_U(A_{1i} \times A_{2i}) \to \Pi_U A_{ji}$, which we will denote $\{\pi_j\}^U$. Here $\{\pi_j\}^U(f^U) = \langle \pi_{ji}(f(i)) : i \in I\rangle^U$ for each $f \in \Pi_I(A_{1i} \times A_{2i})$. The relationship between π_j^U and $\{\pi_j\}^U$ is clarified in Lemma 4.2(1) below.

The relationship between I-tuples of extractions $\varepsilon_{ji} : A_{1i} + A_{2i} \circ\!\!\longrightarrow A_{ji}$ and insertions $\iota_{ji} : A_{ji} \to A_{1i} + A_{2i}$ is preserved by U-liftings: if $x \in \Pi_U(A_{1i} + A_{2i})$ and $y \in \Pi_U A_{ji}$, then $x = \iota_j^U(y)$ iff $x \in \mathrm{Dom}\,\varepsilon_j^U$ and $\varepsilon_j^U(x) = y$.

The U-lifting of an I-tuple of evaluations $eval_i : A_i^D \times D \to A_i$ is the function $eval^U : \Pi_U(A_i^D) \times D^U \to \Pi_U A_i$ having

$$eval^U(f^U, g^U) = \langle eval_i(f(i), g(i))\rangle^U = \langle f(i)(g(i))\rangle^U \tag{4.i}$$

for all $f \in \Pi_I(A_i^D)$ and $g \in \Pi_I D = D^I$. From this it is deducible that for $x \in \Pi_U(A_i^D)$, and $d \in D$,

$$ev_d^U(x) = eval^U(x, \bar{d}^U), \tag{4.ii}$$

where $ev_d^U : \Pi_U(A_i^D) \to \Pi_U A_i$ is the U-lifting of the I-tuple of the functions $ev_{di} : A_i^D \to A_i$ that evaluate at d.

Now fix an I-tuple $\langle A_i \xrightarrow{\alpha_i} |\tau|A_i : i \in I\rangle$ of τ-coalgebras, and let $\alpha = \langle \alpha_i : i \in I\rangle$. The transition structures α_i lift to a function $\alpha^U : \Pi_U A_i \to \Pi_U(|\tau|A_i)$, and the term denotations $[\![\Gamma \rhd M : \sigma]\!]_{\alpha_i}$ lift to a function

$$[\![\Gamma \rhd M : \sigma]\!]^U : \Pi_U A_i \times \Pi_U|\sigma_1|A_i \times \cdots \times \Pi_U|\sigma_n|A_i \longrightarrow \Pi_U|\sigma|A_i \tag{4.iii}$$

where $\sigma_1, \ldots, \sigma_n$ is the list of types of Γ.

α^U is not a τ-coalgebra on A^U since its codomain is $\Pi_U(|\tau|A_i)$ rather than $|\tau|(\Pi_U A_i)$. We wish to define a coalgebraic structure on $\Pi_U A_i$ that interprets terms in a manner related to the functions $[\![\Gamma \rhd M : \sigma]\!]^U$. To achieve this it is necessary to remove some points from $\Pi_U A_i$. The key to understanding which ones are to be retained is provided by considering the U-lifting of the α_i-denotations $[\![M]\!]_{\alpha_i} : A_i \to [\![o]\!]$ of a ground observable term $M : o$. This is the function

$$[\![M]\!]^U = \langle [\![M]\!]_{\alpha_i} : i \in I\rangle^U : \Pi_U A_i \to [\![o]\!]^U.$$

To act as a denotation for M it should assign values in $[\![o]\!]$, viewed as a subset of $[\![o]\!]^U$. In other words we should have

$$[\![M]\!]^U(x) \in e[\![o]\!] = \{\bar{c}^U : c \in [\![o]\!]\} \subseteq [\![o]\!]^U.$$

We are thus led to define an element x of $\Pi_U A_i$ to be *observable* if $[\![M]\!]^U(x) \in e[\![o]\!]$ for every ground observable τ-term $M : o$. If $x = f^U$, this means that for each such M there exists an observable element $c_M \in [\![o]\!]$ such that $[\![M]\!]^U(x) = \bar{c}_M{}^U$ and so

$$\{i \in I : [\![M]\!]_{\alpha_i}(f(i)) = c_M\} \in U. \tag{4.iv}$$

Put $\Pi_U A_i^+ = \{x \in \Pi_U A_i : x \text{ is observable}\}$. The members of $\Pi_U A_i^+$ will be the states of our modified coalgebraic ultraproduct.

In the ultrapower case of $\Pi_U A_i = A^U$, given a single τ-coalgebra (A, α), we will write A^+ for the set $\{x \in A^U : x \text{ is observable}\}$. For each $a \in A$ and any ground $M : o$,

$$[\![M]\!]^U(e_A(a)) = [\![M]\!]^U(\bar{a}^U) = ([\![M]\!]_\alpha \circ \bar{a})^U = \left(\overline{[\![M]\!]_\alpha(a)}\right)^U \in e[\![o]\!],$$

so $e_A(a)$ is observable. Thus e_A embeds A into A^+, allowing us to view A^+ as an extension of A. This also shows that A^+ is non-empty whenever A is non-empty. But it is possible that A^+ is just the e_A-image of A, so there are no "new" observable elements of A^U, even when A is infinite. (If A is infinite, A^U is a proper extension of $e_A(A)$). More generally, in the ultraproduct case it is possible that $\Pi_U A_i^+$ is *empty*, even when all the A_i's are non-empty.

These last claims can be demonstrated by considering coalgebras for the constant functor $\bar{\omega}$, where $\omega = \{0, 1, 2, \dots\}$. For each $n \in \omega$, let (A_n, α_n) be the $\bar{\omega}$-coalgebra having $A_n = \{n, n+1, n+2, \dots\}$ and $\alpha_n : A_n \to \omega$ the inclusion function, i.e. $\alpha_n(x) = x$. Thus if M is the ground observable term $\mathsf{tr}(\mathsf{s})$, then $[\![M]\!]_{\alpha_n}(x) = x$. Let U be a non-principal ultrafilter on ω. Then for any $f \in \Pi_\omega A_n$, and any observable element $c \in \omega$, the set

$$\{n \in \omega : [\![M]\!]_{\alpha_n}(f(n)) = c\} = \{n \in \omega : f(n) = c\}$$

is finite, since in general $f(n) \in A_n$ so $f(n) \geq n$. Hence this set cannot be a member of the non-principal U. This shows that f^U is not an observable element of $\Pi_U A_n$, and hence that $\Pi_U A_n^+$ is empty, even though each A_n is non-empty. Adapting this argument to the single coalgebra (A_0, α_0), for which $A_0 = \omega$ and α_0 is the identity function $\omega \to \omega$, we see that if f^U is an observable element of the ultrapower A_0^U, then there is some $c \in \omega$ such that $\{n \in \omega : f(n) = c\} \in U$ and so $f^U = \bar{c}^U$. So in this case A_0^+ is identifiable with ω and contains no other members of ω^U.

The construction just given can be adapted to build other examples in which $\Pi_U A_n^+$ is non-empty, but has fewer elements than any of the A_n's. We will use this idea again after the proof of Theorem 5.2 to produce a counter example to a certain version of Łoś's Theorem.

At the other extreme, there are types τ for which every ultraproduct of τ-coalgebras is observational, i.e. $\Pi_U A_i^+ = \Pi_U A_i$ always. This happens, for instance, when the denotation $[\![o]\!]$ of any observable subtype of τ is finite. For if $[\![o]\!] = \{c_1, \dots, c_n\}$, then for any element f^U of $\Pi_U A_i$ and any ground term $M : o$, it follows that the finite union

$$\{i \in I : [\![M]\!]_{\alpha_i}(f(i) = c_1\} \cup \dots \cup \{i \in I : [\![M]\!]_{\alpha_i}(f(i) = c_n\},$$

is equal to I, so one of these sets $\{i \in I : [\![M]\!]_{\alpha_i}(f(i) = c_k\}$ belongs to U. This shows that f^U is observable.

Now we return to the general ultraproduct construction and take up the problem of lifting an I-tuple $\alpha = \langle \alpha_i : i \in I \rangle$ of τ-coalgebras to a transition structure of the form

$$\alpha^+ : \Pi_U A_i^+ \to |\tau|(\Pi_U A_i^+).$$

The nature of this transition will depend on the formation of the type τ. Since types are built inductively from their subtypes, the construction of α^+ will involve an induction on type-formation. The notion of a path plays a crucial role, since paths analyse the relationship between a functor and its components. Inevitably we are lead to the following result about the lifting of path actions.

Theorem 4.1 *For any path $|\tau| \overset{p}{\rightsquigarrow} |\sigma|$ beginning at $|\tau|$ there exist partial functions $\langle p_{A_i} \circ \alpha_i \rangle^+ : \Pi_U A_i^+ \circ\!\!\longrightarrow |\sigma| \Pi_U A_i^+$ and $\theta_\sigma : \Pi_U |\sigma| A_i \circ\!\!\longrightarrow |\sigma| \Pi_U A_i^+$,*

$$\begin{array}{ccc}
\Pi_U A_i^+ & \hookrightarrow & \Pi_U A_i \\
\Big\downarrow{\scriptstyle \langle p_{A_i} \circ \alpha_i \rangle^+} & & \Big\downarrow{\scriptstyle \langle p_{A_i} \circ \alpha_i \rangle^U} \\
|\sigma| \Pi_U A_i^+ & \overset{\theta_\sigma}{\twoheadleftarrow\!\circ} & \Pi_U |\sigma| A_i
\end{array}$$

such that θ_σ is surjective (onto $|\sigma| \Pi_U A_i^+$), with

(1) $\mathrm{Dom}\, \langle p_{A_i} \circ \alpha_i \rangle^+ = \Pi_U A_i^+ \cap \mathrm{Dom}\, \langle p_{A_i} \circ \alpha_i \rangle^U$*;*

(2) $x \in \mathrm{Dom}\, \langle p_{A_i} \circ \alpha_i \rangle^+$ *implies* $\langle p_{A_i} \circ \alpha_i \rangle^U(x) \in \mathrm{Dom}\, \theta_\sigma$*;*

and the above diagram commutes wherever defined.

Moreover, in the case of an ultrapower of a single τ-coalgebra (A,α), we have a diagram

$$\begin{array}{ccccc}
A & \xrightarrow{\;e_A\;} & A^+ & \hookrightarrow & A^U \\
\Big\downarrow{\scriptstyle p_A\circ\alpha} & & \Big\downarrow{\scriptstyle (p_A\circ\alpha)^+} & & \Big\downarrow{\scriptstyle (p_A\circ\alpha)^U} \\
[\![\,\sigma\,]\!]_A & \xrightarrow{\;|\sigma|e_A\;} & [\![\,\sigma\,]\!]_{A^+} & \xleftarrow{\;\theta_\sigma\;} & [\![\,\sigma\,]\!]_A^U
\end{array}$$

that commutes whenever defined, such that

(3) $a \in \mathrm{Dom}\,(p_A \circ \alpha)$ *implies* $e_A(a) \in \mathrm{Dom}\,(p_A \circ \alpha)^+$; *and*

(4) $b \in [\![\,\sigma\,]\!]_A$ *implies* $\bar{b}^U \in \mathrm{Dom}\,\theta_\sigma$. □

Note that commuting of the right square of the second diagram of Theorem 4.1 is just the ultrapower case of the commuting of the first diagram, in which $\langle p_{A_i} \circ \alpha_i\rangle^U$ is the U-lifting of the I-tuple $\langle p_{A_i} \circ \alpha_i : i \in I\rangle$.

The details of the proof of the Theorem have been deferred to the final Section 9, along with the proofs of some other technical results to follow. In the present section we focus on explaining the *definitions* of the partial maps $\langle p_{A_i} \circ \alpha_i\rangle^+$ and θ_σ, the latter of which plays an important part in the proof of Łoś's Theorem in Section 5 (see Theorems 5.1 and 5.2).

But first we immediately note that when $\sigma = \tau$ and p is the empty path, so that $p_{A_i} = \mathrm{id}_{A_i}$, the Theorem gives a function $\alpha^+ : \Pi_U A_i^+ \circ\!\!\longrightarrow |\sigma|\Pi_U A_i^+$ whose domain is $\Pi_U A_i^+ \cap \mathrm{Dom}\,\alpha^U = \Pi_U A_i^+$, hence α^+ is total, such that the following diagram commutes.

$$\begin{array}{ccc}
\Pi_U A_i^+ & \hookrightarrow & \Pi_U A_i \\
\Big\downarrow{\scriptstyle \alpha^+} & & \Big\downarrow{\scriptstyle \alpha^U} \\
|\tau|\Pi_U A_i^+ & \xleftarrow{\;\theta_\sigma\;} & \Pi_U|\tau|A_i
\end{array}$$

This gives the definition of α^+ as a τ-coalgebra, which will be called the *observational ultraproduct* of α with respect to U.

Moreover, in the ultrapower case we get the commuting diagram

$$\begin{array}{ccccc}
A & \xrightarrow{\;e_A\;} & A^+ & \hookrightarrow & A^U \\
\Big\downarrow{\scriptstyle \alpha} & & \Big\downarrow{\scriptstyle \alpha^+} & & \Big\downarrow{\scriptstyle \alpha^U} \\
[\![\,\tau\,]\!]_A & \xrightarrow{\;|\tau|e_A\;} & [\![\,\tau\,]\!]_{A^+} & \xleftarrow{\;\theta_\tau\;} & [\![\,\tau\,]\!]_A^U
\end{array}$$

defining α^+ as the *observational ultrapower* of the coalgebra α with respect to U. The left square of this last diagram shows that the injection e_A is a coalgebraic morphism making α isomorphic to a subcoalgebra of α^+.

In proving 4.1 we will need the following results about the extent to which ultraproducts commute with set-theoretic constructions. Parts (1) and (2) of this lemma state that ultraproducts commute with products and coproducts up to bijection, while part (3) gives a weaker statement for powers, involving only a surjection.

Lemma 4.2

(1) *There is a bijection* $\chi_{\mathrm{pr}} : \Pi_U(A_{1i} \times A_{2i}) \to \Pi_U A_{1i} \times \Pi_U A_{2i}$ *defined by*

$$\chi_{\mathrm{pr}}(x) = (\{\pi_1\}^U(x), \{\pi_2\}^U(x)),$$

with $\pi_j^U \circ \chi_{\mathrm{pr}} = \{\pi_j\}^U$.

(2) *There is a bijection* $\chi_{\mathrm{co}} : \Pi_U(A_{1i} + A_{2i}) \to \Pi_U A_{1i} + \Pi_U A_{2i}$ *such that for each* $x \in \mathrm{Dom}\,\chi_{\mathrm{co}}$ *there is exactly one* j *with* $x \in \mathrm{Dom}\,\varepsilon_j^U$ *and* $\chi_{\mathrm{co}}(x) = \iota_j(\varepsilon_j^U(x))$.

(3) *There is a surjection* $\chi_{\mathrm{po}} : \Pi_U(A_i^D) \twoheadrightarrow (\Pi_U A_i)^D$ *defined by* $\chi_{\mathrm{po}}(x)(d) = ev_d^U(x)$.

Proof. See Section 9.1. □

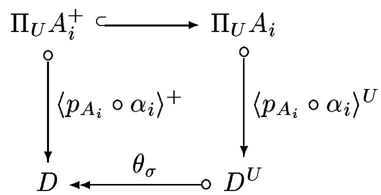
The partial surjections θ_σ of Theorem 4.1 are built using the χ-functions from Lemma 4.2. In the case (3) of powers, χ_{po} will not in general be bijective. For instance, in the ultrapower case we get a surjection $(A^D)^U \twoheadrightarrow (A^U)^D$, but if A is finite then A^U will be of the same size as A, so $(A^U)^D$ will be of the same size as A^D, which may be of strictly lower cardinality than $(A^D)^U$. Moreover, there does not seem to be a naturally definable injection of $(A^U)^D$ into $(A^D)^U$. This accounts for the direction of the θ_σ arrows, which may at first sight appear to be the opposite of what it should be.

The definitions of $\langle p_{A_i} \circ \alpha_i \rangle^+$ and θ_σ proceed by induction on the formation of the type σ, as follows. In each case, $\mathrm{Dom}\,\langle p_{A_i} \circ \alpha_i \rangle^+$ is *defined* by the equation 4.1(1). The justification of the definitions, where required, and the verifications of their properties, may be found in Section 9.2.

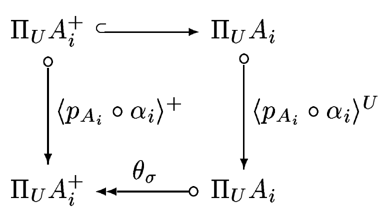
Case $\sigma \in \mathbb{O}$ Here we let $D = [\![\sigma]\!]$, so that $|\sigma|$ is the constant functor $\bar{D}$. Define θ_σ to be the inverse of the injection $e_D : D \rightarrowtail D^U$, and put

$$\langle p_{A_i} \circ \alpha_i \rangle^+(x) = \theta_\sigma(\langle p_{A_i} \circ \alpha_i \rangle^U(x))$$

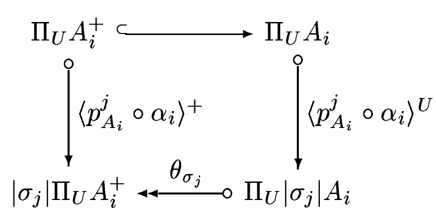
to give the following version of the first diagram of 4.1 in this case.



Case $\sigma = \mathsf{St}$ Here $|\sigma|$ is the identity functor Id. Let θ_σ be the inverse of the inclusion $\Pi_U A_i^+ \hookrightarrow \Pi_U A_i$, and put $\langle p_{A_i} \circ \alpha_i \rangle^+(x) = \langle p_{A_i} \circ \alpha_i \rangle^U(x)$:



Case $\sigma = \sigma_1 \times \sigma_2$ Make the induction hypothesis that the statement of Theorem 4.1 holds for σ_1 and σ_2. From the path $|\tau| \overset{p}{\rightsquigarrow} |\sigma|$ we obtain, for $j = 1$ and 2, the path $p^j = |\tau| \overset{p.\pi_j}{\rightsquigarrow} |\sigma_j|$ and, by the induction hypothesis, a diagram



fulfilling Theorem 4.1, with $p^j_{A_i} = \pi_{ji} \circ p_{A_i}$, where π_{ji} projects $|\sigma_1|A_i \times |\sigma_2|A_i$ onto $|\sigma_j|A_i$.

Let $\theta_{\sigma_1 \times \sigma_2} = (\theta_{\sigma_1} \times \theta_{\sigma_2}) \circ \chi_{\mathrm{pr}}$, the composition of

$$\Pi_U(|\sigma_1|A_i \times |\sigma_2|A_i) \xrightarrow{\chi_{\mathrm{pr}}} \Pi_U|\sigma_1|A_i \times \Pi_U|\sigma_2|A_i \xrightarrow{\theta_{\sigma_1} \times \theta_{\sigma_2}} |\sigma_1|\Pi_U A_i^+ \times |\sigma_2|\Pi_U A_i^+.$$

Define $\langle p_{A_i} \circ \alpha_i\rangle^+$ to be the pairing function $\langle\langle p^1_{A_i} \circ \alpha_i\rangle^+, \langle p^2_{A_i} \circ \alpha_i\rangle^+\rangle$.

$$\begin{array}{ccc} \Pi_U A_i^+ & \hookrightarrow & \Pi_U A_i \\ \downarrow{\scriptstyle \langle p_{A_i}\circ\alpha_i\rangle^+} & & \downarrow{\scriptstyle \langle p_{A_i}\circ\alpha_i\rangle^U} \\ |\sigma_1|\Pi_U A_i^+ \times |\sigma_2|\Pi_U A_i^+ & \xleftarrow{\theta_{\sigma_1\times\sigma_2}} & \Pi_U(|\sigma_1|A_i \times |\sigma_2|A_i) \end{array}$$

Case $\sigma = \sigma_1 + \sigma_2$ Assume the Theorem holds for σ_1 and σ_2. This time we define p^j to be the path $|\tau| \overset{p.\varepsilon_j}{\rightsquigarrow} |\sigma_j|$ and, by the induction hypothesis, have the same diagram

$$\begin{array}{ccc} \Pi_U A_i^+ & \hookrightarrow & \Pi_U A_i \\ \downarrow{\scriptstyle \langle p^j_{A_i}\circ\alpha_i\rangle^+} & & \downarrow{\scriptstyle \langle p^j_{A_i}\circ\alpha_i\rangle^U} \\ |\sigma_j|\Pi_U A_i^+ & \xleftarrow{\theta_{\sigma_j}} & \Pi_U|\sigma_j|A_i \end{array}$$

fulfilling Theorem 4.1, but now $p^j_{A_i} = \varepsilon_{ji} \circ p_{A_i}$.

Let $\theta_{\sigma_1+\sigma_2} = (\theta_{\sigma_1} + \theta_{\sigma_2}) \circ \chi_{\mathrm{co}}$, the composition of

$$\Pi_U(|\sigma_1|A_i + |\sigma_2|A_i) \xrightarrow{\chi_{\mathrm{co}}} \Pi_U|\sigma_1|A_i + \Pi_U|\sigma_2|A_i \xrightarrow{\theta_{\sigma_1} + \theta_{\sigma_2}} |\sigma_1|\Pi_U A_i^+ + |\sigma_2|\Pi_U A_i^+.$$

It then turns out that $\Pi_U A_i^+ \cap \mathrm{Dom}\,\langle p_{A_i} \circ \alpha_i\rangle^U$ is the disjoint union of the domains $\mathrm{Dom}\,\langle p^1_{A_i} \circ \alpha_i\rangle^+$ and $\mathrm{Dom}\,\langle p^2_{A_i} \circ \alpha_i\rangle^+$, so we put

$$\langle p_{A_i} \circ \alpha_i\rangle^+(x) = \iota_j(\langle p^j_{A_i} \circ \alpha_i\rangle^+(x)) \in |\sigma_1|\Pi_U A_i^+ + |\sigma_2|\Pi_U A_i^+$$

for the unique j such that $\langle p^j_{A_i} \circ \alpha_i\rangle^+(x)$ is defined.

$$\begin{array}{ccc} \Pi_U A_i^+ & \hookrightarrow & \Pi_U A_i \\ \downarrow{\scriptstyle \langle p_{A_i}\circ\alpha_i\rangle^+} & & \downarrow{\scriptstyle \langle p_{A_i}\circ\alpha_i\rangle^U} \\ |\sigma_1|\Pi_U A_i^+ + |\sigma_2|\Pi_U A_i^+ & \xleftarrow{\theta_{\sigma_1+\sigma_2}} & \Pi_U(|\sigma_1|A_i + |\sigma_2|A_i) \end{array}$$

Case of $o \Rightarrow \sigma$ Assume the Theorem holds for σ and let $D = [\![o]\!]$. Then from the path $|\tau| \overset{p}{\rightsquigarrow} |o \Rightarrow \sigma|$ we obtain, for each $d \in D$, the path $p^d = |\tau| \overset{p.ev_d}{\rightsquigarrow} |\sigma|$ and, by hypothesis on σ, the diagram

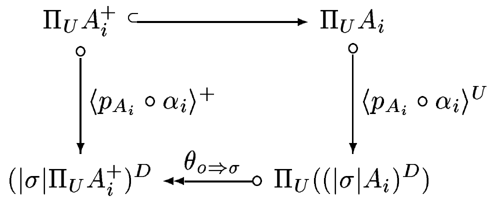
$$\begin{array}{ccc} \Pi_U A_i^+ & \hookrightarrow & \Pi_U A_i \\ \downarrow{\scriptstyle \langle p^d_{A_i}\circ\alpha_i\rangle^+} & & \downarrow{\scriptstyle \langle p^d_{A_i}\circ\alpha_i\rangle^U} \\ |\sigma|\Pi_U A_i^+ & \xleftarrow{\theta_{\sigma}} & \Pi_U|\sigma|A_i \end{array}$$

Let $\theta_{o\Rightarrow\sigma}$ be the composition of

$$\Pi_U((|\sigma|A_i)^D) \xrightarrow{\chi_{\mathrm{po}}} (\Pi_U|\sigma|A_i)^D \xrightarrow{\theta_\sigma^D} (|\sigma|\Pi_U A_i^+)^D.$$

Define $\langle p_{A_i} \circ \alpha_i\rangle^+(x) \in (|\sigma|\Pi_U A_i^+)^D$ by putting

$$\langle p_{A_i} \circ \alpha_i\rangle^+(x)(d) = \langle p^d_{A_i} \circ \alpha_i\rangle^+(x) \in |\sigma|\Pi_U A_i^+.$$



□

That completes the construction of the coalgebra α^+. To conclude this section here are some technical results that will be needed later about how U-lifting preserves the semantics of case terms, applications, and λ-abstractions.

Lemma 4.3 *Suppose Γ is a context with list of types $\tau_1, \ldots, \tau_n$. Let $x \in \Pi_U A_i$ and $\gamma \in \Pi_U |\tau_1| A_i \times \cdots \times \Pi_U |\tau_n| A_i$.*

(1) *Given terms $\Gamma \rhd N : \sigma_1 + \sigma_2$ and $\Gamma, v_j : \sigma_j \rhd M_j : \sigma$ for $j = 1$ and 2, and resultant term $\Gamma \rhd \mathsf{case}(N, M_1, M_2) : \sigma$, then if $[\![\, \Gamma \rhd N \,]\!]^U(x, \gamma) \in \mathrm{Dom}\, \varepsilon_j^U$ we have*

$$[\![\, \Gamma \rhd \mathsf{case}(N, M_1, M_2) \,]\!]^U(x, \gamma) = [\![\, \Gamma, v_j : \sigma_j \rhd M_j \,]\!]^U(x, \gamma, \varepsilon_j^U [\![\, \Gamma \rhd N \,]\!]^U(x, \gamma)).$$

(2) *Given terms $\Gamma \rhd M : o \Rightarrow \sigma$ and $\Gamma \rhd N : o$,*

$$[\![\, \Gamma \rhd M \cdot N : \sigma \,]\!]^U(x, \gamma) = eval^U([\![\, \Gamma \rhd M \,]\!]^U(x, \gamma), [\![\, \Gamma \rhd N \,]\!]^U(x, \gamma)).$$

(3) *Given term $\Gamma, v : o \rhd M : \sigma$, then for any $d \in [\![\, o \,]\!]$,*

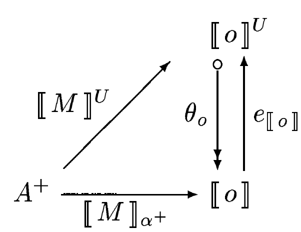
$$ev_d^U([\![\, \Gamma \rhd \lambda v.M : o \Rightarrow \sigma \,]\!]^U(x, \gamma)) = [\![\, \Gamma, v : o \rhd M \,]\!]^U(x, \gamma, \bar{d}^U).$$

Proof. See Section 9.3. □

5 Łoś's Theorem and (Non)Compactness

We are going to eventually show that a polynomial coalgebra (A, α) validates the same observable formulas as any of its observational ultrapowers (Corollary 5.3). To prove this we first have to establish a result (Theorem 5.2), about satisfaction of formulas by elements of an ultraproduct, which is the analogue in this setting of Łoś's Theorem in the standard theory of ultraproducts (see the discussion at the beginning of Section 4).

The use we make of an observational ultrapower α^+ derives ultimately from the fact that for a ground observable term $M : o$, the denotation $[\![\, M \,]\!]_{\alpha^+}$ agrees with the U-lifting $[\![\, M \,]\!]^U$ of $[\![\, M \,]\!]_\alpha : A \to [\![\, o \,]\!]$ in the sense that $[\![\, M \,]\!]_{\alpha^+} = \theta_o \circ ([\![\, M \,]\!]^U \upharpoonright A^+)$, or equivalently $[\![\, M \,]\!]^U \upharpoonright A^+ = e_{[\![\, o \,]\!]} \circ [\![\, M \,]\!]_{\alpha^+}$:



Similarly it can be shown that if M is ground of type St, then $[\![\, M \,]\!]_{\alpha^+}$ is just the restriction of $[\![\, M \,]\!]^U$, i.e. $[\![\, M \,]\!]_{\alpha^+}(x) = [\![\, M \,]\!]^U(x)$ for all $x \in A^+$. But to prove such facts takes an induction on the formation of the ground term $\emptyset \rhd M$, which may involve more complex types and non-empty contexts. Therefore we have to prove a more elaborate result in Theorem 5.1 below.

We formulate this result for any observational ultraproduct $(\Pi_U A_i^+, \alpha^+)$ of an I-tuple of τ-coalgebras (A_i, α_i), with $\alpha = \langle \alpha_i : i \in I \rangle$. Given a context Γ with types $\sigma_1, \ldots, \sigma_n$, let $\theta_\Gamma = \theta_{\sigma_1} \times \cdots \times \theta_{\sigma_n}$ be the product of the functions

$$\theta_{\sigma_k} : \Pi_U |\sigma_k| A_i \circ\!\!\twoheadrightarrow |\sigma_k| \Pi_U A_i^+ = [\![\, \sigma_k \,]\!]_{\alpha^+}$$

given by Theorem 4.1. Then $\mathrm{Dom}\,\theta_\Gamma$ is the product of the $\mathrm{Dom}\,\theta_{\sigma_k}$'s, and so $\Pi_U A_i^+ \times \mathrm{Dom}\,\theta_\Gamma$ is a subset of the domain of the U-lifting $[\![\Gamma \rhd M]\!]^U$ of the denotations $[\![\Gamma \rhd M]\!]_{\alpha_i}$ for any term M in context Γ, since

$$\mathrm{Dom}\,[\![\Gamma \rhd M]\!]^U = \Pi_U A_i \times \Pi_U|\sigma_1|A_i \times \cdots \times \Pi_U|\sigma_n|A_i$$

(see (4.iii)). Note that $\theta_\Gamma : \mathrm{Dom}\,\theta_\Gamma \circ\!\!\twoheadrightarrow [\![\Gamma]\!]_{\alpha^+}$ is surjective, because each θ_{σ_k} is surjective and maps onto $[\![\sigma_k]\!]_{\alpha^+}$.

We can now state the theorem that explains the sense in which $[\![\Gamma \rhd M]\!]_{\alpha^+}$ can be viewed as a restriction of $[\![\Gamma \rhd M]\!]^U$. The main use of this result will be to derive our version of Loś's Theorem in 5.2.

Theorem 5.1 *For any τ-term $\Gamma \rhd M : \sigma$, any $x \in \Pi_U A_i^+$, and any $\gamma \in \mathrm{Dom}\,\theta_\Gamma$:*

$$\begin{array}{ccc} \Pi_U A_i^+ \times \mathrm{Dom}\,\theta_\Gamma & \xrightarrow{[\![\Gamma \rhd M]\!]^U} & \Pi_U|\sigma|A_i \\ \mathrm{id}\times\theta_\Gamma \downarrow & & \downarrow \theta_\sigma \\ \Pi_U A_i^+ \times [\![\Gamma]\!]_{\alpha^+} & \xrightarrow{[\![\Gamma \rhd M]\!]_{\alpha^+}} & |\sigma|\Pi_U A_i^+ \end{array}$$

(1) $[\![\Gamma \rhd M]\!]^U(x,\gamma) \in \mathrm{Dom}\,\theta_\sigma$;

(2) $\theta_\sigma \circ [\![\Gamma \rhd M]\!]^U(x,\gamma) = [\![\Gamma \rhd M]\!]_{\alpha^+}(x, \theta_\Gamma(\gamma))$;

(3) *If* $\sigma \in \mathbb{O}$, $[\![\Gamma \rhd M]\!]^U(x,\gamma) = e_{[\![\sigma]\!]} \circ [\![\Gamma \rhd M]\!]_{\alpha^+}(x, \theta_\Gamma(\gamma))$.

(4) *If* $\sigma = \mathsf{St}$, $[\![\Gamma \rhd M]\!]^U(x,\gamma) = [\![\Gamma \rhd M]\!]_{\alpha^+}(x, \theta_\Gamma(\gamma))$.

Proof. See Section 9.4. □

Theorem 5.2 (Łoś's Theorem for observational ultraproducts)
Let $\Gamma \rhd \varphi$ be an observable τ-formula, with $\Gamma = (v_1 : \sigma_1, \ldots, v_n : \sigma_n)$.

(1) *If $f^U \in \Pi_U A_i^+$, and $g_k^U \in \mathrm{Dom}\,\theta_{\sigma_k}$ for all $k \leq n$, then*

$\alpha^+, f^U, \theta_{\sigma_1}(g_1^U), \ldots, \theta_{\sigma_n}(g_n^U) \models \Gamma \rhd \varphi$ *if, and only if,*

$$\{i \in I : \alpha_i, f(i), g_1(i), \ldots, g_n(i) \models \Gamma \rhd \varphi\} \in U.$$

(2) $\alpha^+, f^U \models \Gamma \rhd \varphi$ *if, and only if,* $\{i \in I : \alpha_i, f(i) \models \Gamma \rhd \varphi\} \in U$.

(3) *If* $\{i \in I : \alpha_i \models \Gamma \rhd \varphi\} \in U$, *then* $\alpha^+ \models \Gamma \rhd \varphi$.

(4) *If every member of $\Pi_U A_i$ is observable, then $\alpha^+ \models \Gamma \rhd \varphi$ implies* $\{i \in I : \alpha_i \models \Gamma \rhd \varphi\} \in U$.

Proof. (1) is proven by induction on the formation of formulas by the rules of Figure 3. The base case is where φ is an equation $M_1 \approx M_2$ with $\Gamma \rhd M_1 : o$ and $\Gamma \rhd M_2 : o$ for some observable type o.

Let $\gamma = (g_1^U, \ldots, g_n^U) \in \mathrm{Dom}\,\theta_\Gamma$ and $\theta_\Gamma(\gamma) = (\theta_{\sigma_1}(g_1^U), \ldots, \theta_{\sigma_n}(g_n^U))$. For $i \in I$, put $g(i) = (g_1(i), \ldots, g_n(i))$. Then

$\alpha^+, f^U, \theta_\Gamma(\gamma) \models \Gamma \rhd M_1 \approx M_2$ iff

$$[\![\Gamma \rhd M_1]\!]_{\alpha^+}(f^U, \theta_\Gamma(\gamma)) = [\![\Gamma \rhd M_2]\!]_{\alpha^+}(f^U, \theta_\Gamma(\gamma)),$$

which, by Theorem 5.1(2) holds iff

$$\theta_o([\![\Gamma \rhd M_1]\!]^U(f^U, \gamma)) = \theta_o([\![\Gamma \rhd M_2]\!]^U(f^U, \gamma)).$$

This in turn is equivalent to

$$[\![\Gamma \rhd M_1]\!]^U(f^U,\gamma) = [\![\Gamma \rhd M_2]\!]^U(f^U,\gamma),$$

because θ_o is injective (being inverse to $[\![o]\!] \rightarrowtail [\![o]\!]^U$). Now this last equation holds iff

$$\langle [\![\Gamma \rhd M_1]\!]_{\alpha_i}(f(i),g(i)) : i \in I\rangle^U = \langle [\![\Gamma \rhd M_2]\!]_{\alpha_i}(f(i),g(i)) : i \in I\rangle^U,$$

which holds iff the set

$$\{i \in I : [\![\Gamma \rhd M_1]\!]_{\alpha_i}(f(i),g(i)) = [\![\Gamma \rhd M_2]\!]_{\alpha_i}(f(i),g(i))\}$$

belongs to U. But this set is just

$$\{i \in I : \alpha_i, f(i), g_1(i), \ldots, g_n(i) \models \Gamma \rhd \varphi\}.$$

Thus (1) holds for all observable equations. The inductive cases then follow by standard arguments from the fact that U is an ultrafilter.

Given (1), to prove (2) let

$$J = \{i \in I : \alpha_i, f(i) \models \Gamma \rhd \varphi\}.$$

Suppose $J \in U$. Then for any $\delta \in [\![\Gamma]\!]_{\alpha^+}$, let $\delta = \theta_\Gamma(\gamma)$ for some $\gamma \in \mathrm{Dom}\,\theta_\Gamma$ (since θ_Γ is surjective). Let $g(i)$ be related to δ as in the proof of (1). Then

$$\{i \in I : \alpha_i, f(i), g(i) \models \Gamma \rhd \varphi\} \supseteq J \in U,$$

so by (1), $\alpha^+, f^U, \delta \models \Gamma \rhd \varphi$. This shows that $\alpha^+, f^U \models \varphi$.

For the converse we reason contrapositively, assuming $J \notin U$. Now for each $i \notin J$ there exist elements $g_k(i) \in [\![\sigma_k]\!]_{A_i}$, for all $k \leq n$, such that

$$\alpha_i, f(i), g_1(i), \ldots, g_n(i) \not\models \Gamma \rhd \varphi.$$

For $i \in J$, choose $g_k(i)$ arbitrarily in $[\![\sigma_k]\!]_{A_i}$. Then the set

$$\{i \in I : \alpha_i, f(i), g_1(i), \ldots, g_n(i) \models \Gamma \rhd \varphi\}$$

is included in J, so if $J \notin U$ then this set cannot belong to U either, hence

$$\alpha^+, f^U, \theta_{\sigma_1}(g_1^U), \ldots, \theta_{\sigma_n}(g_n^U) \not\models \Gamma \rhd \varphi$$

by part (1), showing that $\alpha^+, f^U \not\models \Gamma \rhd \varphi$ and proving (2).

For (3), let $J' = \{i \in I : \alpha_i \models \Gamma \rhd \varphi\}$ and suppose $J' \in U$. Then if $f^U \in \Pi_U A_i^+$, the set $\{i \in I : \alpha_i, f(i) \models \Gamma \rhd \varphi\}$ includes J' and so also belongs to U. Hence by part (2), $\alpha^+, f^U \models \Gamma \rhd \varphi$. Since this holds for all states f^U of α^+, it follows that $\alpha^+ \models \Gamma \rhd \varphi$.

For (4), let $J' = \{i \in I : \alpha_i \models \Gamma \rhd \varphi\}$ as in the proof of (3). Then for each $i \notin J'$ there exist a state $f(i) \in A_i$ and elements $g_k(i)$ as in the proof of part (2), with

$$\alpha_i, f(i), g_1(i), \ldots, g_n(i) \not\models \Gamma \rhd \varphi.$$

For $i \in J'$ we choose $f(i) \in A_i$ and $g_k(i) \in [\![\sigma_k]\!]_{A_i}$ arbitrarily. But now if every member of $\Pi_U A_i$ is observable, then in particular $f^U \in \Pi_U A_i^+$, so we can proceed in the manner of the proof of (2) to argue contrapositively that if $J' \notin U$, then $\alpha^+, f^U, \theta_{\sigma_1}(g_1^U), \ldots, \theta_{\sigma_n}(g_n^U) \not\models \Gamma \rhd \varphi$, and hence $\alpha^+ \not\models \Gamma \rhd \varphi$. □

5.2(4) states that the converse of 5.2(3) holds when $\Pi_U A_i^+ = \Pi_U A_i$ (for example when all sets $[\![o]\!]$ of observable elements are finite, as we saw in Section 4). But this converse of 5.2(3) does not hold in general – by contrast with the satisfaction of first-order formulas in classical ultraproducts. As we saw in Section 4, it is possible for α^+ to be the *empty* coalgebra even when every α_i is non-empty. In that case we have $\alpha^+ \models \bot$ while $\{i \in I : \alpha_i \models \bot\} = \emptyset \notin U$.

To give an example that avoids this rather singular case of the empty coalgebra, we take up again the $\bar{\omega}$-coalgebras (A_n, α_n) from Section 4 that have $A_n = \{n, n+1, \dots\}$ and $\alpha_n : A_n \to \omega$ the inclusion function. Let (Z, ζ) be a fixed $\bar{\omega}$-coalgebra whose state set is finite and disjoint from ω. Define the coalgebra (B_n, β_n) to be the union $(A_n, \alpha_n) + (Z, \zeta)$. Then if U is a non-principal ultrafilter on ω, the resulting observational ultraproduct $(\Pi_\omega B_n^+, \beta^+)$ is just an isomorphic copy of (Z, ζ). For if $f^u \in \Pi_\omega B_n$, then either $\{n : f(n) \in A_n\} \in U$, in which case f^U is not observable by the same argument as for the α_n's, or else $\{n : f(n) \in Z\} \in U$, in which case there is some $z \in Z$ with $f^U = \bar{z}^U$ (by the finiteness of Z). The verification that $\bar{z}^U$ is observable, and that this leads to an isomorphism $\beta^+ \cong \zeta$, is left to the interested reader. (It helps to observe that in general β_n is a subcoalgebra of β_0, from which it follows that for any ground observable term M and any state $x \in B_n$, the value $[\![M]\!]_{\beta_n}(x)$ is independent of n).

Now suppose that Z consists of a single state $\star$ with $\zeta(\star) = 0$. Then the ground observable formula $(\mathsf{tr}(\mathsf{s}) \approx 0)$ valid in ζ, so $\beta^+ \models (\mathsf{tr}(\mathsf{s}) \approx 0)$. But each β_n has states x with $\beta_n(x) \neq 0$, so $\{n \in \omega : \beta_n \models (\mathsf{tr}(\mathsf{s}) \approx 0)\} = \emptyset \notin U$, giving another counter-example to the converse of 5.2(3).

On the other hand the converse of 5.2(3) does always hold for ultrapowers. Indeed, the class of all models of an observable formula is invariant under observational ultrapowers:

Corollary 5.3 *Let α^+ be an observational ultrapower of a τ-coalgebra α over an ultrafilter U. If $\Gamma \rhd \varphi$ is an observable τ-formula, then*

$$\alpha \models \Gamma \rhd \varphi \text{ if, and only if, } \alpha^+ \models \Gamma \rhd \varphi.$$

Proof. The implication from left to right is given by 5.2(3) for this case: if $\alpha_i = \alpha$ for all $i \in I$, and $\alpha \models \Gamma \rhd \varphi$, then

$$\{i \in I : \alpha_i \models \Gamma \rhd \varphi\} = I \in U,$$

so by 5.2(3), $\alpha^+ \models \Gamma \rhd \varphi$.

For the converse, let $\alpha^+ \models \Gamma \rhd \varphi$. Then for any $a \in A$, $\alpha^+, \bar{a}^U \models \Gamma \rhd \varphi$, so by 5.2(2) the set

$$\{i \in I : \alpha, \bar{a}(i) \models \Gamma \rhd \varphi\}$$

belongs to U. Hence this set is non-empty, implying that $\alpha, a \models \Gamma \rhd \varphi$. This shows that $\alpha \models \Gamma \rhd \varphi$.

(Note that this converse follows alternatively from Theorem 2.2, which states that the class of models of an observable formula is closed under subcoalgebras and isomorphism. Here α is isomorphic to a subcoalgebra of α^+ under the embedding $e_A : A \to A^+$.) □

Compactness

One of the fundamental uses of ultraproducts and Łoś's Theorem in first-order logic is to prove the Compactness Theorem, stating that a set of sentences must have a model if each of its finite subsets has a model (see [6, p. 102]). In the present coalgebraic setting this statement is trivially fulfilled, because every set of formulas whatsoever has a model – the empty coalgebra. To avoid this vacuity we will say that *the Compactness Theorem holds for the type τ* if the following is true:

> a set of observable τ-formulas has a *non-empty* model whenever each of its finite subsets has a non-empty model.

Another approach is to consider satisfaction of formulas at particular states, rather than validity in models. A set Φ of formulas will be called *satisfiable* if there exists a coalgebra α and a state x of α such that every member of Φ is true at x in α. Then an alternative formulation of compactness is the statement:

> a set of observable τ-formulas is satisfiable whenever each of its finite subsets is satisfiable.

It turns out that both of these formulations fail for coalgebraic logic. This can be seen from our earlier example involving the $\bar{\omega}$-coalgebras (A_n, α_n) with $A_n = \{n, n+1, \dots\}$ and α_n the inclusion function $A_n \hookrightarrow \omega$. Let φ_n be the (ground observable) formula $(\mathsf{tr}(\mathsf{s}) \not\approx n)$ and $\Phi_\omega = \{\varphi_n : n \in \omega\}$. At any state x of a $\bar{\omega}$-coalgebra α there will be some n such that φ_n is false at x, namely $n = \alpha(x)$. Hence Φ_ω is not satisfiable, and so has no non-empty models. But each of its finite subsets is satisfiable, and indeed has a non-empty model, because

$$\alpha_n \models \varphi_0 \wedge \cdots \wedge \varphi_{n-1}.$$

Compactness does hold for some types, such as those for which every ultraproduct is observational (e.g. those for which each $[\![o]\!]$ is finite). More generally we have

Theorem 5.4 *Suppose that every observational ultraproduct of non-empty τ-coalgebras is non-empty. Then the Compactness Theorem holds for τ.*

Proof. In this case the standard ultraproduct proof of Compactness applies, as we now confirm. Let Φ be a set of observable τ-formulas, and take I to be the set of all finite subsets of Φ. Suppose that each $i \in I$ has a non-empty model α_i. For each formula $\varphi \in \Phi$, let $I_\varphi = \{i \in I : \varphi \in i\}$. Then $I_\varphi \subseteq \{i \in I : \alpha_i \models \varphi\}$, because $\alpha_i \models i$.

Now the collection $\{I_\varphi : \varphi \in \Phi\}$ has the finite intersection property, because

$$\{\varphi_1, \dots, \varphi_n\} \in I_{\varphi_1} \cap \cdots \cap I_{\varphi_n}.$$

Hence there is an ultrafilter U on I extending $\{I_\varphi : \varphi \in \Phi\}$. Then for each $\varphi \in \Phi$, the set $\{i \in I : \alpha_i \models \varphi\}$ belongs to U, so $\alpha^+ \models \varphi$ by Łoś's Theorem 5.2(3).

This shows that α^+ is a model of Φ. But by hypothesis, α^+ is non-empty.

□

6 Enlarging Ultrapowers

A set Φ of ground formulas is *satisfiable in coalgebra* α if there is some state of α at which all members of Φ are true, i.e. some $x \in A$ such that $\alpha, x \models \varphi$ for all $\varphi \in \Phi$. Φ is *finitely satisfiable in* α if each finite subset of Φ is satisfiable in α (different finite subsets of Φ may be satisfied at different states of α). Putting $\varphi^\alpha = \{x \in A : \alpha, x \models \varphi\}$, we see that Φ is finitely satisfiable in α iff the collection $\Phi^\alpha = \{\varphi^\alpha : \varphi \in \Phi\}$ of subsets of A has the *finite intersection property*, i.e. every non-empty finite subset of Φ^α has non-empty intersection.

There is a well-known construction in the theory of ultrapowers that will enable us to force certain finitely α-satisfiable Φ's to become satisfiable in some observational ultrapower α^+ of α. By choosing a suitable ultrafilter U it can be arranged that

> *any collection $\mathcal{S}$ of subsets of A with the finite intersection property has a "non-standard element in its intersection". This element is an $f^U \in A^U$ such that for each $C \in \mathcal{S}$, $f^U \in_U C$, which means that $\{i : f(i) \in C\} \in U$.*

If this property holds, then α^+ will be called an *enlarging* observational ultrapower of α.

To see how the enlarging property can be enforced, let I_A be the set of all finite subsets of the powerset of A. A typical element of I_A is of the form $i = \{C_1, \dots, C_n\}$ with the C_j's being subsets of A. For each $k \in I_A$, let $I_k = \{i \in I_A : k \subseteq i\}$. The collection $U_A = \{I_k : k \in I_A\}$ has the finite intersection property, since $I_{k_1} \cap \cdots \cap I_{k_n}$ contains the element $i = k_1 \cup \cdots \cup k_n$. Let U be any ultrafilter on I_A that extends U_A.

Now if $\mathcal{S}$ is a collection of subsets of A with the finite intersection property, let $f : I_A \to A$ be any function such that $f(i) \in \bigcap(i \cap \mathcal{S})$ whenever $i \cap \mathcal{S} \neq \emptyset$. Note that by the finite intersection property, if $i \cap \mathcal{S} \neq \emptyset$ then $\bigcap(i \cap \mathcal{S}) \neq \emptyset$, so such an f does exist. Then for any $C \in \mathcal{S}$, put $k = \{C\} \in I_A$. Observe that if $i \in I_{\{C\}}$ then $C \in i \cap \mathcal{S}$, so $f(i) \in C$. This shows that $I_{\{C\}} \subseteq \{i : f(i) \in C\}$, and therefore $f^U \in_U C$. Hence f^U is in the intersection of $\mathcal{S}$, as desired.

Thus we see that any coalgebra does indeed have enlarging observational ultrapowers. Here now is our main application of this construction.

Theorem 6.1 *Let τ be a type that has at least one non-trivial observable subtype. Let α and β be τ-coalgebras, and α^+ an enlarging observational ultrapower of α. Then the bisimilarity relation from α^+ to β is surjective if, and only if, every ground observable τ-formula valid in α is valid also in β.*

Proof. Surjectivity of bisimilarity here means that for each state of y of β there is a state x of α with $x \sim y$. If this holds then every observable formula valid in α^+ is valid in β, because validity of observable formulas is preserved by surjective bisimulations (see Theorem 2.2. The details are in [10, Theorem 5.4]). Then any observable formula valid in α is also valid in α^+ by Corollary 5.3, hence is valid in β.

For the converse, assume that every ground observable formula valid in α is valid in β. Let y be any state of β. If $M : o$ is any ground observable term, let $c_M = [\![M]\!]_\beta(y) \in [\![o]\!]$. Let Φ_y be the set of equations $M \approx c_M$ for all ground observable M. By definition, Φ_y is satisfied by y in β.

Each finite $\{\varphi_1, \dots, \varphi_n\} \subseteq \Phi_y$ is satisfiable in α, for otherwise the ground observable formula

$$\neg(\varphi_1 \wedge \cdots \wedge \varphi_n)$$

would be valid in α, hence valid in β by hypothesis, contrary to the fact that this formula is false at y. This establishes that Φ_y is finitely satisfiable in α, which means, as noted in the first paragraph of this section, that the collection $\Phi_y^\alpha = \{\varphi^\alpha : \varphi \in \Phi_y\}$ of subsets of A has the finite intersection property.

It follows by the *enlarging* property that if U is the ultrafilter that gives rise to α^+, then there is some $f^U \in A^U$ such that for each equation $(M \approx c_M) \in \Phi_y$ we have $f^U \in_U (M \approx c_M)^\alpha$, which means that the set

$$\begin{aligned} I_M &= \{i \in I : \alpha, f(i) \models M \approx c_M\} \\ &= \{i \in I : [\![M]\!]_\alpha(f(i)) = c_M\} \end{aligned}$$

belongs to U. Since this holds for all ground observable M, f^U is observable by (4.iv), so $f^U \in A^+$. Also, since $I_M \in U$, Łoś's Theorem 5.2(2) implies that $\alpha^+, f^U \models M \approx c_M$, so

$$[\![M]\!]_{\alpha^+}(f^U) = c_M = [\![M]\!]_\beta(y).$$

Therefore f^U and y assign the same values to all ground observable terms, and so are bisimilar by Theorem 3.2(4). This establishes that each state of β is bisimilar to a state in A^+, as desired. □

Surjectivity of bisimilarity from α^+ to β can be understood as meaning that all possible behaviours of β are represented in α^+, since each state of β is bisimilar to one of α^+. Theorem 6.1 states that this situation obtains precisely when β is a model of what we may call the *observational theory* of α^+, the set of all ground observable formulas that are valid in α^+.

Let us call two coalgebras α and β *totally bisimilar* if each state of one is bisimilar to a state of the other, i.e. bisimilarity is surjective both from α to β and from β to α. This means that α and β are behaviourally equivalent: any behaviour of either is represented in the other. Other equivalent formulations of this notion are that bisimilarity from α to β is a total and surjective relation, and that the bisimilarity relations in both directions are total.

Two coalgebras will be called *logically equivalent* if they have the same logically expressible properties, which we define to mean that they validate the same ground observable formulas, i.e. they have the same observational theory. Totally bisimilar coalgebras must be logically equivalent, by the preservation of validity under surjective bisimulations. Here is an exact relationship between these two notions of equivalence of coalgebras:

Theorem 6.2 *Two coalgebras are logically equivalent if, and only if, they have totally bisimilar observational ultrapowers.*

Proof. Suppose α and β have observational ultrapowers α^+ and β^+ that are totally bisimilar. Then α^+ and β^+ are logically equivalent. But by Corollary 5.3, α is logically equivalent to α^+, and β to β^+. Hence α and β are logically equivalent.

Conversely, if α and β are logically equivalent, take *enlarging* observational ultrapowers α^+ and β^+. Since β is logically equivalent to β^+ (Corollary 5.3), β^+ is then a model of the observational theory of α. Hence by Theorem 6.1, with β^+ in place of β, it follows that bisimilarity is surjective from α^+ to β^+.

Interchanging the roles of α and β in this argument yields that α^+ and β^+ are totally bisimilar. □

7 Logically Definable Classes of Coalgebras

The machinery needed to give a structural characterisation of logically definable classes of coalgebras has now all been exposed. The following result is the analogue for polynomial functors of Theorem 9.2 of [11] for monomial functors.

Theorem 7.1 *If τ has at least one non-trivial observable subtype, then for any class K of τ-coalgebras, the following are equivalent.*

(1) *K is the class of all models of some set of ground observable formulas.*

(2) *K is the class of all models of some set of rigid observable formulas.*

(3) *K is closed under disjoint unions, images of bisimulations, and observational ultrapowers.*

(4) *K is closed under disjoint unions, images of bisimilarity relations, and enlarging observational ultrapowers.*

Proof. (1) implies (2) by definition. Theorem 2.2 and Corollary 5.3 explain why (2) implies (3). (3) implies (4) by definition. Thus it remains to show that (4) implies (1).

Let Φ be the set of all ground observable formulas that are valid in all members of K. By definition all members of K are models of Φ, so it suffices to prove that all models of Φ are members of K to establish (1). Let β be any model of Φ. For each ground observable φ such that $\beta \not\models \varphi$ there must be some $\alpha_\varphi \in K$ such that $\alpha_\varphi \not\models \varphi$, or else φ is valid in all members of K, hence φ belongs to Φ, hence $\beta \models \varphi$ contrary to assumption. Let α be the disjoint union of all these α_φ's from K. Since α_φ is isomorphic to a subcoalgebra of α, any observable formula valid in α will be valid in α_φ. Thus $\alpha \not\models \varphi$ when $\beta \not\models \varphi$. In other words, any ground observable formula valid in α is valid in every α_φ, hence valid in β. Therefore if we take α^+ to be an enlarging observational ultrapower of α, then by Theorem 6.1 the bisimilarity relation from α^+ to β is surjective.

In sum: β is the image under bisimilarity of an enlarging observational ultrapower of a disjoint union of coalgebras from K. The closure conditions listed in (4) then ensure that $\beta \in K$. □

There are a number of equivalent ways of expressing the closure conditions identified in Theorem 7.1. A *covariety* is a class K that is closed under subcoalgebras, images of coalgebraic morphisms, and disjoint unions; while a *behavioural* covariety is one that is also closed under images of total bisimulations (the origin of this terminology is explained in Section 8). In fact a behavioural covariety K is closed under images of arbitrary bisimulations, not just total ones, for if R is a surjective bisimulation from α to β, then the domain $\operatorname{Dom} R$ is a subcoalgebra of α and R is a total bisimulation from $\operatorname{Dom} \alpha$ onto β, so $\alpha \in K$ implies $\beta \in K$ by closure under subcoalgebras and images of total bisimulations.

On the other hand, a class of coalgebras that is closed under images of bisimulations must be closed under domains and images of morphisms (the domain of a morphism is the image of the inverse bisimulation of that morphism). Hence the class is closed under subcoalgebras, since a subcoalgebra is the domain of the inclusion morphism.

Altogether then we see that a behavioural covariety is precisely the same thing as a class K that is closed under images of bisimulations and disjoint unions. So according to Theorem 7.1, the model classes of sets of rigid observable formulas are precisely the *behavioural covarieties that are closed under observational ultrapowers.*

Recall that $Mod\,\Phi$ is the class of all models of a set Φ of formulas, and $Mod\,\varphi$ is the class of all models of a single formula φ. Using the notion of observational ultrapower, we can now give structural conditions under which a class of coalgebras is of the form $Mod\,\varphi$:

Theorem 7.2 *Let K be a class of τ-coalgebras, where τ has at least one non-trivial observable subtype. If K is closed under disjoint unions, images of bisimulations, and observational ultrapowers, and the complement of K is closed under observational ultraproducts, then K is the class of all models of a single ground observable formula.*

Proof. By the complement of K we mean the class of all τ-coalgebras that are not in K.

Now by Theorem 7.1 and the given closure conditions on K we know that $K = Mod\,\Phi$ for some set Φ of ground observable formulas. Let I be the set of all finite subsets of Φ. We will show that $K = Mod\,i$ for some $i \in I$.

Suppose this does not hold. Then for each $i \in I$, $K \neq Mod\,i$, so as $K \subseteq Mod\,i$ there must be some τ-coalgebra α_i with $\alpha_i \in Mod\,i$, i.e. $\alpha_i \models i$, but $\alpha_i \notin K$. By the construction in the compactness proof of Theorem 5.4, there is then an ultrafilter U on I such that the resulting observational ultraproduct α^+ of the α_i's over U is a model of Φ. Hence $\alpha^+ \in K$. But this contradicts the assumption that the complement of K is closed under observational ultraproducts.

It follows that we must have $K = Mod\,\{\varphi_1, \ldots, \varphi_n\}$, where $\varphi_1, \ldots, \varphi_n$ are some finite number of members of Φ. Then $K = Mod\,\varphi$ where φ is the single ground observable formula $\varphi_1 \wedge \cdots \wedge \varphi_n$. □

The converse of Theorem 7.2 requires that if φ is ground and observable then the complement of $Mod\,\varphi$ is closed under observational ultraproducts. This is true when all ultraproducts of τ-coalgebras are observational (by 5.2(4)), but it is not true in general. As we have seen, there are examples where $\alpha_i \not\models \varphi$ for all $i \in I$, but $\alpha^+ \models \varphi$.

To conclude this discussion of logically definable classes of coalgebras, we exhibit a class of the form $Mod\,\Phi$ that is not equal to $Mod\,\varphi$ for any observable formula φ. Take Φ to be the set $\Phi_\omega = \{\varphi_n : n \geq 1\}$ of τ-formulas from the game machine example discussed at the end of Section 2. Recall the collection $\{\alpha_n : n \geq 1\}$ of τ-coalgebras constructed there such that $\alpha_n \models \varphi_k$ whenever $k < n$, but $\alpha_n \not\models \varphi_n$. Hence $\alpha_n \notin Mod\,\Phi_\omega$ for all n.

Now let U be a non-principal ultrafilter on $\{n : n \geq 1\}$, and α^+ the observational ultraproduct of the α_n's over U. Then for each $k \geq 1$, the set $\{n : \alpha_n \models \varphi_k\}$ includes the cofinite set $\{n : n > k\}$, and so belongs to U. Hence by 5.2(3), $\alpha^+ \models \varphi_k$. Thus $\alpha^+ \in Mod\,\Phi_\omega$, showing that the complement of $Mod\,\Phi_\omega$ is not closed under observational ultraproducts.

The type τ in this example involves the observable types data, 1 and bool. If we make the (reasonable) assumption that $[\![\mathsf{data}]\!]$ is finite, then all three sets of observable elements are finite, and so all ultraproducts of τ-coalgebras are observational, i.e. have $\Pi_U A_i^+ = \Pi_U A_i$. Hence by 5.2(4) the complement of $Mod\,\varphi$ is closed under observational ultraproducts for any observable τ-formula φ. This demonstrates that $Mod\,\Phi_\omega$ is not equal to $Mod\,\varphi$ for any such φ.

8 The Analogy With Birkhoff's Theorem

Many coalgebraic concepts arise by categorical duality from concepts from the theory of algebras. The nature of this duality is that a notion defined by some diagram with arrows gives rise to a dual notion by reversing all the arrows. Thus coproducts are dual to products, monomorphisms to epimorphisms, hence subobjects to epimorphic images and vice versa, etc.

In suggesting that Theorem 7.1 is an analogue of the theorem of Birkhoff [7] we are not claiming that 7.1 is the *dual* of Birkhoff's result. Rather, it is being viewed as result of the same species, in a way that we now explain.

Birkhoff's theorem states that a certain *logical* description of a class of abstract algebras is equivalent to another *structural* characterization. The logical description is that of an *equationally definable* class: the class of all models of some set of equations. Such a class is sometimes called a *variety*, borrowing a term from classical algebraic geometry, where a variety is (or was originally) the set of solutions of some system of equations[1]. The structural characterization is that of closure under homomorphic images (H), subalgebras (S) and direct products (P). Birkhoff showed that a variety is the same thing as an HSP-closed class, and this is now sometimes called the Variety Theorem, or the HSP Theorem.

Subsequently there were developed many results in classical model theory stating that a class of structures is the class of all models of a set of sentences having a particular syntactic form if, and only if, it is closed under certain algebraic constructions. The deepest result of this kind (deepest in terms of difficulty of proof) is the theorem of Keisler and Shelah that a class K is *basic elementary*, i.e. is the class of models of some first-order sentence, iff K and its complement are both closed under isomorphism and ultraproducts. Many results in this area are expressed as *preservation theorems*, stating that a sentence or formula is logically equivalent to one of a special syntactic form iff its validity is preserved by certain constructions. Thus a theorem of Łoś and Tarski states that a first-order sentence is equivalent to a universal sentence (one with only universal quantifiers in its prenex normal form) iff it is preserved by substructures. Chang, Łoś and Suszko characterized the sentences preserved by unions of chains as those equivalent to universal-existential sentences. Lyndon characterized syntactically the sentences preserved under homomorphisms, and those preserved by subdirect products, while Keisler characterized those preserved by reduced products. These and other such results about inverse limits, direct products et alia are discussed extensively in [9, Sections 3.2, 6.2, 6.3], [8, Section V.2] and [12, Chapter 7].

Theorems 7.1 and 7.2 of this paper may be seen as results that belong to this model-theoretic tradition. We have developed a syntax and semantics that provide a natural logic for polynomial coalgebras, analogous to the natural role of equations in abstract algebra. This view is supported by the role played by observable terms and formulas in logically specifying such coalgebraic concepts as bisimilarity, morphism, the action of path functions, and modal assertions about state transitions [10]. We have then sought to find appropriate structural closure conditions that identify the model classes defined by observable formulas. This has required, not just a dualization of the HSP constructions, but the introduction of the new construction of observational ultrapowers.

Over time Birkhoff's theorem turned into a definition: it became common to *define* a variety as being an HSP-closed class. Then as coalgebraic theory developed, the term *covariety* was adopted [35, 36] for the dual notion of a class of coalgebras closed under subcoalgebras, images of morphisms and disjoint unions (coproducts) . A number of papers developed approaches to characterizing such covarieties by concepts replacing Birkhoff's equational satisfaction. Rutten [35, 36, Section 17] worked with the idea of a *colouring* $A \to C$ of the state-set A of a coalgebra, and showed that for a given covariety K there is a "colour set" C and a special subcoalgebra α_K of the coalgebra α_C cofreely generated over C, such that each C-colouring of a coalgebra $\alpha \in K$ lifts to a morphism $\alpha \to \alpha_C$ that factors through α_K. This lifting property characterizes the members of K, and moreover, any subcoalgebra of α_C determines a covariety in this way. These results were established for coalgebras of endofunctors on **Set** that have a "boundedness" condition guaranteeing the existence of cofree coalgebras.

The notion of a colouring dualizes that of a *variable assignment* $X \to \mathcal{A}$, giving values to a set X of variables in an algebra $\mathcal{A}$. Rutten's use of a subcoalgebra of a cofree coalgebra dualizes the situation of a set E of equations in the variables X determining a quotient $\mathcal{F}_X \twoheadrightarrow \mathcal{F}_E$ of the free algebra $\mathcal{F}_X$ of terms in X, this quotient being given by the smallest congruence on $\mathcal{F}_X$ containing the pairs of terms of the equations in E. An algebra $\mathcal{A}$ is a model of E iff each assignment $X \to \mathcal{A}$ lifts to a homomorphism $\mathcal{F}_X \to \mathcal{A}$ that factors through $\mathcal{F}_X \twoheadrightarrow \mathcal{F}_E$.

The importance of subcoalgebras of cofree coalgebras had already been demonstrated by Jacobs [22] who dualized the notion of a congruence on an algebra to that of a *mongruence*, a subset of a coalgebra that is closed under the coalgebraic operations. He showed how a set E of coalgebraic equations determines a subcoalgebra of a cofree coalgebra, based on the

[1] I have been told by Bernhard Neumann and Paul Cohn that Philip Hall used the term "variety" for equationally defined classes of algebras in lectures in the 1940's.

largest mongruence satisfying E. This construction was used to prove that for a polynomial $T : \mathbf{Set} \to \mathbf{Set}$, the forgetful functor on the category of T-coalgebras satisfying E has a right adjoint and is comonadic. Jacobs concluded by asking if there was a "Birkhoff variety" theorem in this context.

Gumm and Schröder [15] analysed Rutten's construction further by studying, for β a subcoalgebra of α, the class $Q(\alpha, \beta)$ of coalgebras ρ with the property that any morphism $\rho \to \alpha$ factors through β. They showed that $Q(\alpha, \beta)$ is a covariety if α has the *extension property* that given any monomorphism $m : \alpha \rightarrowtail \alpha'$, each morphism with domain α factors through m. Rutten's work showed that for coalgebras of a bounded functor, every covariety is of the form $Q(\alpha_C, \beta)$ with α_C cofree over C. Gumm and Schröder proved that in the bounded case, a coalgebra has the extension property iff it is a retract of a cofree algebra. They also studied covarieties that are closed under images of *total* bisimulations, showing that they can be characterized as the classes $Q(\alpha_C, \beta)$ where α_C is a *final* coalgebra, which means that it is cofree over a *one-element* colour set C. They called such covarieties *complete*, while [3] later suggested the more evocative name *behavioural*. As noted earlier, the model classes of Theorem 7.1 are precisely those behavioural covarieties that are closed under observational ultrapowers.

Gumm [13, 14] developed the idea of a *coequation with covariables in* X as an element of a coalgebra cofreely generated over X, viewing such an element as a "behaviour pattern with variables in X". Satisfaction of a coequation by a coalgebra α meant that it was omitted from the image of any morphism from α to the cofree coalgebra, so that the coequation constituted a "forbidden behaviour pattern". Covarieties then prove to be the same as classes of models of sets of such coequations. A coequation over a single colour/covariable would appear to be an abstract analogue of the formulas in this paper that have a single state parameter s.

Kurz [27, 29] formulated a covariety theorem by developing an abstract version of the notion of a modal logic, including abstract definitions of formulas, models and satisfaction. Roşu [33] gave a characterization of equationally specifiable classes of coalgebras that is closer to the spirit of the present paper, involving closure under certain constructions, one of which is defined in terms of satisfaction of equations, so is logical rather than structural.

An early attempt to "categorize" the notion of an equation was that of Hatcher [16], who worked with the idea of an identity/equation in an arbitrary category as being a pair of arrows with the same domain and codomain, abstracting from the perception of an equation as a pair of members of a free algebra of terms. He characterised categories closed under products and subobjects in terms of "satisfaction" of such identities. This work was generalized by Herrlich, Ringel and Banaschewski [20, 5]. The notion of an equation with variables in X as a pair t, u of arrows into the free algebra $\mathcal{F}X$ generated by X became the notion of an epimorphism $e : \mathcal{F}X \to \mathcal{E}$ (the coequaliser of t and u). An algebra $\mathcal{A}$ satisfies the "equation" e if every morphism $\mathcal{F}X \to \mathcal{A}$ factors through e. In category-theoretic language, this says that $\mathcal{A}$ is *injective* with respect to e. It corresponds to the idea that $\mathcal{A}$ satisfies the equation $t \approx u$ when any function $X \to \mathcal{A}$, lifts to a morphism $\mathcal{F}X \to \mathcal{A}$ that identifies t and u. From this emerged the conception of an *equational* category as one whose objects are just those that are injective with respect to a class of epimorphisms of a certain kind. Birkhoff-style theorems then state that such equational categories can be characterized by categorical closure conditions. Dualizations of Birkhoff's theorem along these lines has been extensively developed by Awodey and Hughes [3, 21, 4], Kurz [28] and Adámek and Porst [2] for coalgebras of endfunctors on a wide range of categories defined by categorical axioms. These results take the form of showing that a covariety is *coequational* in the sense that it consists of those coalgebras that are *pro*-jective with respect to some class of *mono*-morphisms. In many cases these monomorphisms, or *coequations*, correspond to subobjects of cofree coalgebras.

9 Proofs of Technical Results

9.1 Proof of Lemma 4.2

(1) For $f \in \Pi_I(A_{1i} \times A_{2i})$ we get $\pi_{ji} \circ f(i) \in A_{ji}$ where $\pi_{ji} : A_{1i} \times A_{2i} \to A_{ji}$ is the projection. Thus we can define

$$\chi_{\mathrm{pr}}(f^U) = (\langle \pi_{1i} \circ f(i) \rangle^U, \langle \pi_{2i} \circ f(i) \rangle^U) = (\{\pi_1\}^U(f^U), \{\pi_2\}^U(f^U)).$$

The proof that this gives a well-defined bijection is routine. Since χ_{pr} is just the pairing function $\langle \{\pi_1\}^U, \{\pi_2\}^U \rangle$ and π_j^U is the j-th projection from $\Pi_U A_{1i} \times \Pi_U A_{2i}$ it is immediate that $\pi_j^U \circ \chi_{\mathrm{pr}} = \{\pi_j\}^U$.

(2) For $f \in \Pi_I(A_{1i} + A_{2i})$, the sets $\{i \in I : f(i) \in \iota_{1i} A_{1i}\}$ and $\{i \in I : f(i) \in \iota_{2i} A_{2i}\}$ are disjoint and have union I, so exactly one of them belongs to the ultrafilter U. In other words, there is exactly one $j = 1$ or 2 for which $\{i \in I : f(i) \in \iota_{ji} A_{ji}\} \in U$. Since $\iota_{ji} A_{ji} = \mathrm{Dom}\, \varepsilon_{ji}$, it follows that $f^U \in \mathrm{Dom}\, \varepsilon_j^U$, where $\varepsilon_j^U : \Pi_U(A_{1i} + A_{2i}) \circ\!\!\longrightarrow \Pi_U A_{ji}$ is the U-lifting of the extractions ε_{ji}. For this j put

$$\chi_{\mathrm{co}}(f^U) = \iota_j(\varepsilon_j^U(f^U)) \in \Pi_U A_{1i} + \Pi_U A_{2i}.$$

Again it is straightforward to show that χ_{co} is bijective.

(3) Recall that $ev_d^U : \Pi_U(A_i^D) \to \Pi_U A_i$ is the U-lifting of the functions $ev_{di} : A_i^D \to A_i$ that evaluate at d.

If $x \in \Pi_U(A_i^D)$, then the function $d \mapsto ev_d^U(x)$ belongs to $(\Pi_U A_i)^D$, and we take it to be $\chi_{\mathrm{po}}(x)$. To see that χ_{po} as defined is surjective, take any $g \in (\Pi_U A_i)^D$. Then for each $d \in D$, $g(d)$ is equal to g_d^U for some $g_d \in \Pi_I A_i$. Now define $f \in \Pi_I(A_i^D)$ by the formula $f(i)(d) = g_d(i)$. Thus $g_d(i) = ev_{di}(f(i))$, so

$$\chi_{\mathrm{po}}(f^U)(d) = ev_d^U(f^U) = \langle ev_{di}(f(i)) : i \in I \rangle^U = g_d^U = g(d),$$

and hence $\chi_{\mathrm{po}}(f^U) = g$. □

9.2 Proof of Theorem 4.1

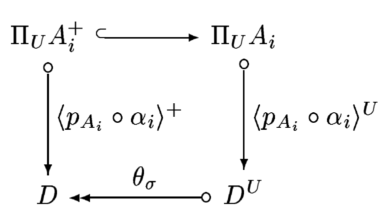
For a given τ, the proof proceeds by induction on the formation of the end-type σ of the path p. The definitions of $\langle p_{A_i} \circ \alpha_i \rangle^+$ and θ_σ are repeated here. In each case of the proof we verify 4.1(1)–(4) and show that the first diagram of 4.1 commutes when defined. For the second diagram we then have only to check that the left square commutes when defined, since the right square is just the ultrapower case of the first diagram.

Let $M_p = \bar{p}[\mathsf{tr}(\mathsf{s})/v]$ be the ground term of type σ given by the Path Lemma 3.1. Then the function $p_{A_i} \circ \alpha_i$ agrees with $[\![M_p]\!]_{\alpha_i}$ whenever it is defined, which implies that

$$\langle p_{A_i} \circ \alpha_i \rangle^U(x) = [\![M_p]\!]^U(x) \text{ whenever } x \in \mathrm{Dom}\, \langle p_{A_i} \circ \alpha_i \rangle^U.$$

Case $\sigma \in \mathbb{O}$ Here we have $D = [\![\sigma]\!]$, $|\sigma|$ is the constant functor $\bar{D}$, and θ_σ is the inverse of the injection $e_D : D \rightarrowtail D^U$. Since e_D is a total function, this makes θ_σ an injective partial mapping *onto* $D = [\![\sigma]\!]_{A^+}$, so θ_σ is surjective.

Now for each $x \in \Pi_U A_i^+ \cap \mathrm{Dom}\, \langle p_{A_i} \circ \alpha_i \rangle^U$, from above $\langle p_{A_i} \circ \alpha_i \rangle^U(x) = [\![M_p]\!]^U(x)$, which belongs to $e_D(D)$ since M_p is of observable type and $x \in \Pi_U A_i^+$, i.e. x is an observable element. But $e_D(D) = \mathrm{Dom}\, \theta_\sigma$, so we can *define* $\mathrm{Dom}\, \langle p_{A_i} \circ \alpha_i \rangle^+$ to be $\Pi_U A_i^+ \cap \mathrm{Dom}\, \langle p_{A_i} \circ \alpha_i \rangle^U$ and put $\langle p_{A_i} \circ \alpha_i \rangle^+(x) = \theta_\sigma(\langle p_{A_i} \circ \alpha_i \rangle^U(x))$ to make 4.1(1) and 4.1(2) true and the following diagram commute as required.



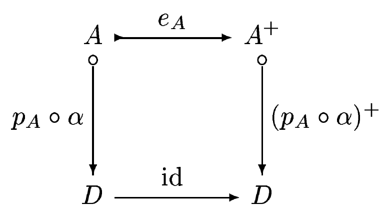
Moreover, in the ultrapower case, if $a \in \mathrm{Dom}\,(p_A \circ \alpha)$, then $\bar{a} \in_U \mathrm{Dom}\,(p_A \circ \alpha)$, so

$$e_A(a) = \bar{a}^U \in \mathrm{Dom}\,(p_A \circ \alpha)^U \cap A^+ = \mathrm{Dom}\,(p_A \circ \alpha)^+,$$

proving 4.1(3). Moreover

$$(p_A \circ \alpha)^U(\bar{a}^U) = \overline{p_A \circ \alpha(a)}^U = e_A(p_A \circ \alpha(a)),$$

so $(p_A \circ \alpha)^+(e_A(a)) = \theta_\sigma((p_A \circ \alpha)^U(\bar{a}^U)) = \theta_\sigma(e_A(p_A \circ \alpha(a))) = p_A \circ \alpha(a)$, showing that the diagram



commutes, which is the left square of the second diagram of 4.1 in this case.

Finally, for 4.1(4), if $b \in [\![\sigma]\!]$, then $\bar{b}^U = e_D(b) \in \mathrm{Dom}\,\theta_\sigma$ as required.

Case $\sigma = \mathsf{St}$ Here $|\sigma|$ is the identity functor Id and θ_σ is the inverse of the inclusion $\Pi_U A_i^+ \hookrightarrow \Pi_U A_i$. Hence θ_σ maps onto $\Pi_U A_i^+ = |\sigma|\Pi_U A_i^+$, so is surjective.

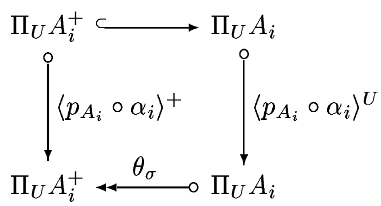
For each $x \in \Pi_U A_i^+ \cap \mathrm{Dom}\,\langle p_{A_i} \circ \alpha_i\rangle^U$, we have $\langle p_{A_i} \circ \alpha_i\rangle^U(x) \in \Pi_U A_i^+$. To see this, let $N : o$ be any ground observable term. Then by the rule (s-Subst), $N[M_p/\mathsf{s}] : o$ is a ground observable term, and so $[\![N[M_p/\mathsf{s}]]\!]^U(x) \in e[\![o]\!]$ as x is observable. Then

$[\![N]\!]^U(\langle p_{A_i} \circ \alpha_i\rangle^U(x)) = [\![N]\!]^U([\![M_p]\!]^U(x)) = \langle [\![N]\!]_{\alpha_i} \circ [\![M_p]\!]_{\alpha_i}\rangle^U(x) =$
(by the semantics of s-Subst) $\langle [\![N[M_p/\mathsf{s}]]\!]_{\alpha_i}\rangle^U(x) = [\![N[M_p/\mathsf{s}]]\!]^U(x)$

which belongs to $e[\![o]\!]$. Since this holds for all ground observable N, it follows that $\langle p_{A_i} \circ \alpha_i\rangle^U(x)$ is observable as desired. Thus we can define

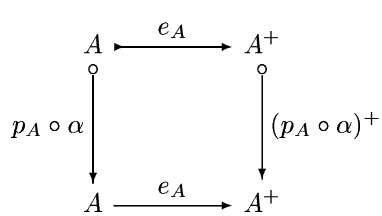
$$\mathrm{Dom}\,\langle p_{A_i} \circ \alpha_i\rangle^+ = \Pi_U A_i^+ \cap \mathrm{Dom}\,\langle p_{A_i} \circ \alpha_i\rangle^U$$

and $\langle p_{A_i} \circ \alpha_i\rangle^+(x) = \langle p_{A_i} \circ \alpha_i\rangle^U(x)$ to make



commute and 4.1(1) and 4.1(2) be true in this case.

For the ultrapower situation, it follows just as in the previous case $\sigma \in \mathbb{O}$ that if $a \in \mathrm{Dom}\,(p_A \circ \alpha)$, then $e_A(a) \in \mathrm{Dom}\,(p_A \circ \alpha)^+$, giving 4.1(3), and $(p_A \circ \alpha)^+(e_A(a)) = \theta_\sigma(e_A(p_A \circ \alpha(a)))$, which in this case is just $e_A(p_A \circ \alpha(a))$, so



commutes, which is the left square of the second diagram of 4.1 when $\sigma = \mathsf{St}$.

For 4.1(4), if $b \in [\![\mathsf{St}]\!]_A = A$, $\bar{b}^U \in A^+ = \mathrm{Dom}\,\theta_{\mathsf{St}}$.

Case $\sigma = \sigma_1 \times \sigma_2$ In this first inductive case we make the induction hypothesis that the statement of Theorem 4.1 holds for σ_1 and σ_2. From the path $|\tau|\overset{p}{\rightsquigarrow}|\sigma|$ we obtain, for $j = 1$ and 2, the path $p^j = |\tau|\overset{p.\pi_j}{\rightsquigarrow}|\sigma_j|$ and, by the induction hypothesis, the diagram

$$\begin{array}{ccc}
\Pi_U A_i^+ & \hookrightarrow & \Pi_U A_i \\
\downarrow{\scriptstyle \langle p^j_{A_i} \circ \alpha_i\rangle^+} & & \downarrow{\scriptstyle \langle p^j_{A_i} \circ \alpha_i\rangle^U} \\
|\sigma_j|\Pi_U A_i^+ & \overset{\theta_{\sigma_j}}{\twoheadleftarrow} & \Pi_U|\sigma_j|A_i
\end{array}$$

fulfilling Theorem 4.1. Here $p^j_{A_i} = \pi_{ji} \circ p_{A_i}$, where π_{ji} projects $|\sigma_1|A_i \times |\sigma_2|A_i$ onto $|\sigma_j|A_i$.

Let $\theta_{\sigma_1\times\sigma_2} = (\theta_{\sigma_1} \times \theta_{\sigma_2}) \circ \chi_{\text{pr}}$, the composition of

$$\Pi_U(|\sigma_1|A_i \times |\sigma_2|A_i) \xrightarrow{\chi_{\text{pr}}} \Pi_U|\sigma_1|A_i \times \Pi_U|\sigma_2|A_i \xrightarrow{\theta_{\sigma_1} \times \theta_{\sigma_2}} |\sigma_1|\Pi_U A_i^+ \times |\sigma_2|\Pi_U A_i^+.$$

The induction hypothesis gives that θ_{σ_1} and θ_{σ_2} are surjective, from which it follows readily that $\theta_{\sigma_1} \times \theta_{\sigma_2}$ is surjective. Since χ_{pr} is a bijection (Lemma 4.2) it then follows that $\theta_{\sigma_1\times\sigma_2} = \chi_{\text{pr}} \circ (\theta_{\sigma_1} \times \theta_{\sigma_2})$ is surjective.

From the definition of χ_{pr} as $\langle\{\pi_1\}^U, \{\pi_2\}^U\rangle$ we see that in general

$$x \in \text{Dom}\,\theta_{\sigma_1\times\sigma_2} \quad \text{iff} \quad \{\pi_j\}^U(x) \in \text{Dom}\,\theta_{\sigma_j} \text{ for } j = 1 \text{ and } 2. \tag{9.i}$$

If $x \in \text{Dom}\,\theta_{\sigma_1\times\sigma_2}$ then $\theta_{\sigma_1\times\sigma_2}(x) = (\theta_{\sigma_1}(\{\pi_1\}^U(x)), \theta_{\sigma_2}(\{\pi_2\}^U(x)))$ and

$$\pi_j(\theta_{\sigma_1\times\sigma_2}(x)) = \theta_{\sigma_j}(\{\pi_j\}^U(x)). \tag{9.ii}$$

Now as π_{ji} and α_i are total, $\text{Dom}\,(p^j_{A_i} \circ \alpha_i) = \text{Dom}\,(p_{A_i} \circ \alpha_i)$. Hence

$$\text{Dom}\,\langle p^j_{A_i} \circ \alpha_i\rangle^+ = \Pi_U A_i^+ \cap \text{Dom}\,\langle p^j_{A_i} \circ \alpha_i\rangle^U = \Pi_U A_i^+ \cap \text{Dom}\,\langle p_{A_i} \circ \alpha_i\rangle^U$$

independently of j. Thus justifies our defining $\langle p_{A_i} \circ \alpha_i\rangle^+$ to be the pairing function $\langle\langle p^1_{A_i} \circ \alpha_i\rangle^+, \langle p^2_{A_i} \circ \alpha_i\rangle^+\rangle$ on $\Pi_U A_i^+ \cap \text{Dom}\,\langle p_{A_i} \circ \alpha_i\rangle^U$,

$$\begin{array}{ccc}
\Pi_U A_i^+ & \hookrightarrow & \Pi_U A_i \\
\downarrow{\scriptstyle \langle p_{A_i} \circ \alpha_i\rangle^+} & & \downarrow{\scriptstyle \langle p_{A_i} \circ \alpha_i\rangle^U} \\
|\sigma_1|\Pi_U A_i^+ \times |\sigma_2|\Pi_U A_i^+ & \overset{\theta_{\sigma_1\times\sigma_2}}{\twoheadleftarrow} & \Pi_U(|\sigma_1|A_i \times |\sigma_2|A_i)
\end{array}$$

and thereby satisfying 4.1(1).

For 4.1(2), let $x \in \text{Dom}\,\langle p_{A_i} \circ \alpha_i\rangle^+ = \text{Dom}\,\langle p^j_{A_i} \circ \alpha_i\rangle^+$. Then for $j = 1, 2$, $\{\pi_j\}^U(\langle p_{A_i} \circ \alpha_i\rangle^U(x)) = \langle p^j_{A_i} \circ \alpha_i\rangle^U(x)$, which belongs to $\text{Dom}\,\theta_{\sigma_j}$ by 4.1(2) for σ_j. Hence by condition (9.i), $\langle p_{A_i} \circ \alpha_i\rangle^U(x) \in \text{Dom}\,\theta_{\sigma_1\times\sigma_2}$, as required. Moreover in this case,

$$\begin{array}{lll}
\pi_j[\theta_{\sigma_1\times\sigma_2}(\langle p_{A_i} \circ \alpha_i\rangle^U(x))] & = \theta_{\sigma_j}(\{\pi_j\}^U(\langle p_{A_i} \circ \alpha_i\rangle^U(x))) & \text{by (9.ii)} \\
& = \theta_{\sigma_j}(\langle p^j_{A_i} \circ \alpha_i\rangle^U(x)) & \text{from above} \\
& = \langle p^j_{A_i} \circ \alpha_i\rangle^+(x) & \text{diagram for } \sigma_j \\
& = \pi_j[\langle p_{A_i} \circ \alpha_i\rangle^+(x)] &
\end{array}$$

for $j = 1$ and 2. Hence $\theta_{\sigma_1\times\sigma_2}(\langle p_{A_i} \circ \alpha_i\rangle^U(x)) = \langle p_{A_i} \circ \alpha_i\rangle^+(x)$, making the last diagram commute.

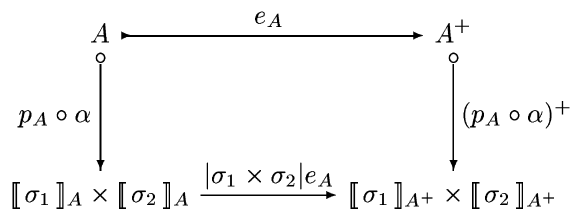
For the ultrapower case, the hypothesis gives commuting of

$$\begin{array}{ccc}
A & \overset{e_A}{\rightarrowtail} & A^+ \\
\downarrow{\scriptstyle p^j_A \circ \alpha} & & \downarrow{\scriptstyle (p^j_A \circ \alpha)^+} \\
[\![\sigma_j]\!]_A & \xrightarrow{|\sigma_j|e_A} & [\![\sigma_j]\!]_{A^+}
\end{array}$$

For 4.1(3), let $a \in \mathrm{Dom}\,(p_A \circ \alpha) = \mathrm{Dom}\,(p^j_A \circ \alpha)$. Then by 4.1(3) for σ_j, $e_A(a) \in \mathrm{Dom}\,(p^j_A \circ \alpha)^+ = \mathrm{Dom}\,(p_A \circ \alpha)^+$ as required. Also

$$\begin{aligned}
\pi_j[(p_A \circ \alpha)^+(e_A(a))] &= (p^j_A \circ \alpha)^+(e_A(a)) \\
&= |\sigma_j| e_A(p^j_A \circ \alpha(a)) && \text{last diagram} \\
&= |\sigma_j| e_A(\pi_j(p_A \circ \alpha(a))) \\
&= \pi_j[\,|\sigma_1 \times \sigma_2| e_A(p_A \circ \alpha(a))], && \text{definition of } |\sigma_1 \times \sigma_2|
\end{aligned}$$

for $j = 1$ and 2, so $(p_A \circ \alpha)^+(e_A(a)) = |\sigma_1 \times \sigma_2| e_A(p_A \circ \alpha(a))$ and the diagram

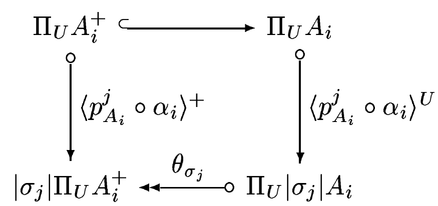


commutes as well.

For 4.1(4), let $b \in [\![\sigma_1]\!]_A \times [\![\sigma_2]\!]_A$. For each j, $\pi_j(b) \in [\![\sigma_j]\!]_A$. But $\{\pi_j\}^U(\bar{b}^U) = (\pi_j \circ \bar{b})^U = \left(\overline{\pi_j(b)}\right)^U$, which belongs to $\mathrm{Dom}\,\theta_{\sigma_j}$ by 4.1(4) assumed inductively for σ_j. Hence by (9.i), $\bar{b}^U \in \mathrm{Dom}\,\theta_{\sigma_1 \times \sigma_2}$.

This completes the inductive proof that Theorem 4.1 holds for $\sigma_1 \times \sigma_2$.

Case $\sigma = \sigma_1 + \sigma_2$ Assume the Theorem holds for σ_1 and σ_2. This time we define p^j to be the path $|\tau| \overset{p.\varepsilon_j}{\rightsquigarrow} |\sigma_j|$ and, by the induction hypothesis, have the same diagram



fulfilling Theorem 4.1, but now $p^j_{A_i} = \varepsilon_{ji} \circ p_{A_i}$, where ε_{ji} is the (partial) extraction from $|\sigma_1|A_i + |\sigma_2|A_i$ to $|\sigma_j|A_i$.

Let $\theta_{\sigma_1+\sigma_2} = (\theta_{\sigma_1} + \theta_{\sigma_2}) \circ \chi_{\mathrm{co}}$, the composition of

$$\Pi_U(|\sigma_1|A_i + |\sigma_2|A_i) \xrightarrow{\chi_{\mathrm{co}}} \Pi_U|\sigma_1|A_i + \Pi_U|\sigma_2|A_i \xrightarrow{\theta_{\sigma_1} + \theta_{\sigma_2}} |\sigma_1|\Pi_U A_i^+ + |\sigma_2|\Pi_U A_i^+.$$

Since θ_{σ_1} and θ_{σ_2} are surjective by induction hypothesis, then so is $\theta_{\sigma_1} + \theta_{\sigma_2}$, and hence so is $\theta_{\sigma_1+\sigma_2}$ as χ_{co} is bijective.

If $x \in \Pi_U(|\sigma_1|A_i + |\sigma_2|A_i)$, then we have $x \in \mathrm{Dom}\,\theta_{\sigma_1+\sigma_2}$ iff $\chi_{\mathrm{co}}(x) \in \mathrm{Dom}\,(\theta_{\sigma_1} + \theta_{\sigma_2})$. But by Lemma 4.2(2) there is a unique j with $x \in \mathrm{Dom}\,\varepsilon^U_j$ and $\chi_{\mathrm{co}}(x) = \iota_j(\varepsilon^U_j(x))$, where $\varepsilon_j = \langle \varepsilon_{ji} \rangle$. We then see that $x \in \mathrm{Dom}\,\theta_{\sigma_1+\sigma_2}$ iff $\varepsilon^U_j(x) \in \mathrm{Dom}\,\theta_{\sigma_j}$, with

$$\theta_{\sigma_1+\sigma_2}(x) = \iota_j(\theta_{\sigma_j}(\varepsilon^U_j(x))).$$

Thus we have

$$x \in \mathrm{Dom}\,\theta_{\sigma_1+\sigma_2} \text{ iff } \varepsilon^U_j(x)\!\downarrow \text{ and } \varepsilon^U_j(x) \in \mathrm{Dom}\,\theta_{\sigma_j} \text{ for some } j. \tag{9.iii}$$

In view of the relationship between extractions and insertions we can also conclude that

$$x \in \mathrm{Dom}\,\theta_{\sigma_1+\sigma_2} \text{ iff } x = \iota^U_j(y) \text{ for some } j \text{ and some } y \in \mathrm{Dom}\,\theta_{\sigma_j}. \tag{9.iv}$$

Hence for any $y \in \mathrm{Dom}\,\theta_{\sigma_j}$, by putting $x = \iota^U_j(y)$ in the above we get $y = \varepsilon^U_j(x)$ and so

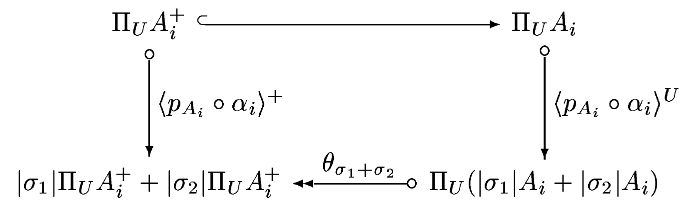
$$\theta_{\sigma_1+\sigma_2}(\iota^U_j(y)) = \iota_j(\theta_{\sigma_j}(y)). \tag{9.v}$$

Now as $p^j_{A_i} = \varepsilon_{ji} \circ p_{A_i}$, $a \in \mathrm{Dom}\, p^j_{A_i} \circ \alpha_i$ iff $a \in \mathrm{Dom}\, p_{A_i} \circ \alpha_i$ and $p_{A_i} \circ \alpha_i(a) \in \mathrm{Dom}\, \varepsilon_{ji}$. Since $|\sigma_1|A_i + |\sigma_2|A_i$ is the disjoint union of $\mathrm{Dom}\, \varepsilon_{1i}$ and $\mathrm{Dom}\, \varepsilon_{2i}$, $\mathrm{Dom}\, (p_{A_i} \circ \alpha_i)$ is the disjoint union of $\mathrm{Dom}\, (p^1_{A_i} \circ \alpha_i)$ and $\mathrm{Dom}\, (p^2_{A_i} \circ \alpha_i)$. From this it can be seen that $\Pi_U A_i^+ \cap \mathrm{Dom}\, \langle p_{A_i} \circ \alpha_i \rangle^U$ is the disjoint union of $\Pi_U A_i^+ \cap \mathrm{Dom}\, \langle p^1_{A_i} \circ \alpha_i \rangle^U$ and $\Pi_U A_i^+ \cap \mathrm{Dom}\, \langle p^2_{A_i} \circ \alpha_i \rangle^U$, and hence is the disjoint union of $\mathrm{Dom}\, \langle p^1_{A_i} \circ \alpha_i \rangle^+$ and $\mathrm{Dom}\, \langle p^2_{A_i} \circ \alpha_i \rangle^+$ by 4.1(1) for σ_1 and σ_2. Moreover, $x \in \mathrm{Dom}\, \langle p^j_{A_i} \circ \alpha_i \rangle^U$ iff $x \in \mathrm{Dom}\, \langle p_{A_i} \circ \alpha_i \rangle^U$ and $\langle p_{A_i} \circ \alpha_i \rangle^U(x) \in \mathrm{Dom}\, \varepsilon^U_j$.

Thus we can define $\mathrm{Dom}\, \langle p_{A_i} \circ \alpha_i \rangle^+ = \Pi_U A_i^+ \cap \mathrm{Dom}\, \langle p_{A_i} \circ \alpha_i \rangle^U$ as in previous cases, and put

$$\langle p_{A_i} \circ \alpha_i \rangle^+(x) = \iota_j(\langle p^j_{A_i} \circ \alpha_i \rangle^+(x)) \in |\sigma_1|\Pi_U A_i^+ + |\sigma_2|\Pi_U A_i^+$$

for the unique j such that $\langle p^j_{A_i} \circ \alpha_i \rangle^+(x)$ is defined.

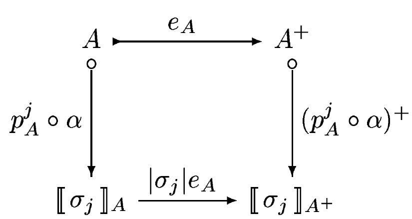


Now when $\langle p^j_{A_i} \circ \alpha_i \rangle^+(x)$ is defined, so too are $\langle p^j_{A_i} \circ \alpha_i \rangle^U(x)$ and $\langle p_{A_i} \circ \alpha_i \rangle^U(x)$, with $\langle p_{A_i} \circ \alpha_i \rangle^U(x) \in \mathrm{Dom}\, \varepsilon^U_j$. But since $\varepsilon_{ji} \circ p_{A_i} = p^j_{A_i}$, we see that $\varepsilon^U_j(\langle p_{A_i} \circ \alpha_i \rangle^U(x)) = \langle p^j_{A_i} \circ \alpha_i \rangle^U(x)$, which belongs to $\mathrm{Dom}\, \theta_{\sigma_j}$ by 4.1(2) for σ_j. From (9.iii) it follows that $\langle p_{A_i} \circ \alpha_i \rangle^U(x) \in \mathrm{Dom}\, \theta_{\sigma_1+\sigma_2}$, fulfilling 4.1(2) for $\sigma_1 + \sigma_2$. Also

$$\begin{aligned}
\theta_{\sigma_1+\sigma_2}(\langle p_{A_i} \circ \alpha_i \rangle^U(x)) &= \iota_j(\theta_{\sigma_j}(\varepsilon^U_j(\langle p_{A_i} \circ \alpha_i \rangle^U(x)))) && \text{definition of } \theta_{\sigma_1+\sigma_2} \\
&= \iota_j(\theta_{\sigma_j}(\langle p^j_{A_i} \circ \alpha_i \rangle^U(x))) && \text{from above} \\
&= \iota_j(\langle p^j_{A_i} \circ \alpha_i \rangle^+(x)) && \text{diagram for } \sigma_j \\
&= \langle p_{A_i} \circ \alpha_i \rangle^+(x),
\end{aligned}$$

so the last diagram commutes as required.

For the ultrapower case, again the hypothesis gives commuting of



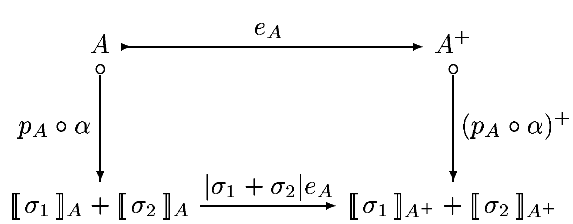
Now let $a \in \mathrm{Dom}\, (p_A \circ \alpha)$. Then $a \in \mathrm{Dom}\, (p^j_A \circ \alpha)$ for a unique j, and so by 4.1(3) for σ_j,

$$e_A(a) \in \mathrm{Dom}\, (p^j_A \circ \alpha)^+ \subseteq \mathrm{Dom}\, (p_A \circ \alpha)^+,$$

proving 4.1(3) for $\sigma_1 + \sigma_2$. Also

$$\begin{aligned}
(p_A \circ \alpha)^+(e_A(a)) &= \iota_j((p^j_A \circ \alpha)^+(e_A(a))) && \text{by definition} \\
&= \iota_j(|\sigma_j|e_A(p^j_A \circ \alpha(a))) && \text{last diagram} \\
&= \iota_j(|\sigma_j|e_A(\varepsilon_j(p_A \circ \alpha(a)))) && \\
&= |\sigma_1 + \sigma_2|e_A(p_A \circ \alpha(a)), && \text{definition of } |\sigma_1 + \sigma_2|,
\end{aligned}$$

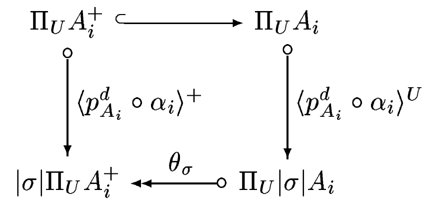
so the the following diagram commutes as required:



For 4.1(4), let $b \in [\![\sigma_1]\!]_A + [\![\sigma_2]\!]_A$. Then $b = \iota_j(c)$ for some j and some $c \in [\![\sigma_j]\!]_A$. By 4.1(4) for σ_j, $\bar{c}^U \in \mathrm{Dom}\,\theta_{\sigma_j}$. But $\bar{b}^U = \iota_j^U(\bar{c}^U)$, so by (9.iv), $\bar{b}^U \in \mathrm{Dom}\,\theta_{\sigma_1+\sigma_2}$.

Hence Theorem 4.1 holds for $\sigma_1 + \sigma_2$.

Case of $o \Rightarrow \sigma$ Assume the Theorem holds for σ and let $D = [\![o]\!]$. Then from the path $|\tau| \overset{p}{\rightsquigarrow} |o \Rightarrow \sigma|$ we obtain, for each $d \in D$, the path $p^d = |\tau| \overset{p.ev_d}{\rightsquigarrow} |\sigma|$ and, by hypothesis on σ, the diagram



fulfilling Theorem 4.1. Here $p^d_{A_i} = ev_{di} \circ p_{A_i}$, with $ev_{di} : (|\sigma|A_i)^D \to |\sigma|A_i$, so as ev_{di} and α_i are total, $\mathrm{Dom}\,(p^d_{A_i} \circ \alpha_i) = \mathrm{Dom}\,(p_{A_i} \circ \alpha_i)$.

Let $\theta_{o\Rightarrow\sigma}$ be the composition of

$$\Pi_U((|\sigma|A_i)^D) \xrightarrow{\chi_{\mathrm{po}}} (\Pi_U|\sigma|A_i)^D \overset{\theta^D_\sigma}{\circ\!\!\longrightarrow} (|\sigma|\Pi_U A_i^+)^D,$$

where the partial function θ^D_σ has $g \in \mathrm{Dom}\,\theta^D_\sigma$ iff for all $d \in D$, $g(d) \in \mathrm{Dom}\,\theta_\sigma$; and $\theta^D_\sigma(g) = \theta_\sigma \circ g$. Thus if $\theta^D_\sigma(g)$ is defined, $\theta^D_\sigma(g)(d) = \theta_\sigma(g(d))$ for all $d \in D$.

Since θ_σ is surjective by induction hypothesis, given any $h \in (|\sigma|\Pi_U A_i^+)^D$ we can *define* a g by the condition that $\theta_\sigma(g(d)) = h(d)$ for all d. Then $\theta^D_\sigma(g) = h$, so θ^D_σ is surjective, and therefore so is $\theta_{o\Rightarrow\sigma}$, since χ_{po} is surjective by Lemma 4.2(3).

Now in general $x \in \mathrm{Dom}\,\theta_{o\Rightarrow\sigma}$ iff $\chi_{\mathrm{po}}(x) \in \mathrm{Dom}\,\theta^D_\sigma$. Since $\chi_{\mathrm{po}}(x)(d) = ev^U_d(x)$, it follows that

$$x \in \mathrm{Dom}\,\theta_{o\Rightarrow\sigma} \quad \text{iff} \quad \text{for all } d \in [\![o]\!],\ ev^U_d(x) \in \mathrm{Dom}\,\theta_\sigma; \tag{9.vi}$$

and when $\theta_{o\Rightarrow\sigma}(x)$ is defined,

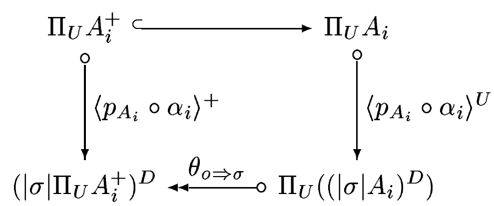
$$\theta_{o\Rightarrow\sigma}(x)(d) = \theta_\sigma(ev^U_d(x)). \tag{9.vii}$$

Now since $\mathrm{Dom}\,(p^d_{A_i} \circ \alpha_i) = \mathrm{Dom}\,(p_{A_i} \circ \alpha_i)$, by 4.1(1) for σ,

$$\mathrm{Dom}\,\langle p^d_{A_i} \circ \alpha_i\rangle^+ = \Pi_U A_i^+ \cap \mathrm{Dom}\,\langle p^d_{A_i} \circ \alpha_i\rangle^U = \Pi_U A_i^+ \cap \mathrm{Dom}\,\langle p_{A_i} \circ \alpha_i\rangle^U$$

for all $d \in D$. So we define $\mathrm{Dom}\,\langle p_{A_i} \circ \alpha_i\rangle^+ = \Pi_U A_i^+ \cap \mathrm{Dom}\,\langle p_{A_i} \circ \alpha_i\rangle^U$ as usual, and specify $\langle p_{A_i} \circ \alpha_i\rangle^+(x) \in (|\sigma|\Pi_U A_i^+)^D$ by putting

$$\langle p_{A_i} \circ \alpha_i\rangle^+(x)(d) = \langle p^d_{A_i} \circ \alpha_i\rangle^+(x) \in |\sigma|\Pi_U A_i^+.$$



For 4.1(2), let $x \in \mathrm{Dom}\,\langle p_{A_i} \circ \alpha_i\rangle^+$. Then for any $d \in D$, $x \in \mathrm{Dom}\,\langle p^d_{A_i} \circ \alpha_i\rangle^+$ and

$$ev^U_d(\langle p_{A_i} \circ \alpha_i\rangle^U(x)) = \langle ev_{di} \circ p_{A_i} \circ \alpha_i\rangle^U(x) = \langle p^d_{A_i} \circ \alpha_i\rangle^U(x),$$

which belongs to $\mathrm{Dom}\,\theta_\sigma$ by 4.1(2) for σ. It follows by condition (9.vi) that $\langle p_{A_i} \circ \alpha_i\rangle^U(x) \in \mathrm{Dom}\,\theta_{o\Rightarrow\sigma}$, as required. Moreover in this case,

$$\begin{aligned}
\theta_{o\Rightarrow\sigma}(\langle p_{A_i} \circ \alpha_i\rangle^U(x))(d) &= \theta_\sigma(ev^U_d(\langle p_{A_i} \circ \alpha_i\rangle^U(x))) && \text{by (9.vii)}\\
&= \theta_\sigma(\langle p^d_{A_i} \circ \alpha_i\rangle^U(x)) && \text{from above}\\
&= \langle p^d_{A_i} \circ \alpha_i\rangle^+(x) && \text{diagram for } \sigma\\
&= \langle p_{A_i} \circ \alpha_i\rangle^+(x)(d).
\end{aligned}$$

This shows that $\theta_{o \Rightarrow \sigma}(\langle p_{A_i} \circ \alpha_i \rangle^U(x)) = \langle p_{A_i} \circ \alpha_i \rangle^+(x)$, making the right square of the last diagram commute.

For the ultrapower case, for each $d \in D$ the induction hypothesis on σ gives the commuting of the diagram

$$\begin{array}{ccc} A & \xrightarrow{\;e_A\;} & A^+ \\ {\scriptstyle p_A^d \circ \alpha}\big\downarrow & & \big\downarrow{\scriptstyle (p_A^d \circ \alpha)^+} \\ [\![\, \sigma \,]\!]_A & \xrightarrow{\;|\sigma| e_A\;} & [\![\, \sigma \,]\!]_{A^+} \end{array}$$

For 4.1(3), let $a \in \mathrm{Dom}\,(p_A \circ \alpha)$. Then for any $d \in D$, $a \in \mathrm{Dom}\,(p_A^d \circ \alpha)$ so by 4.1(3) for σ, $e_A(a) \in \mathrm{Dom}\,(p_A^d \circ \alpha)^+ = \mathrm{Dom}\,(p_A \circ \alpha)^+$ as required. Also

$$\begin{array}{lll} (p_A \circ \alpha)^+(e_A(a))(d) & = (p_A^d \circ \alpha)^+(e_A(a)) & \\ & = |\sigma| e_A(p_A^d \circ \alpha(a)) & \text{last diagram} \\ & = |\sigma| e_A(ev_d(p_A \circ \alpha(a))) & \\ & = |\sigma| e_A((p_A \circ \alpha(a))(d)) & \\ & = |o \Rightarrow \sigma| e_A(p_A \circ \alpha(a))(d) & \text{definition of } |o \Rightarrow \sigma| \end{array}$$

so $(p_A \circ \alpha)^+(e_A(a)) = |o \Rightarrow \sigma| e_A(p_A \circ \alpha(a))$ and the diagram

$$\begin{array}{ccc} A & \xrightarrow{\;e_A\;} & A^+ \\ {\scriptstyle p_A \circ \alpha}\big\downarrow & & \big\downarrow{\scriptstyle (p_A \circ \alpha)^+} \\ [\![\, \sigma \,]\!]_A^D & \xrightarrow{\;|o \Rightarrow \sigma| e_A\;} & [\![\, \sigma \,]\!]_{A^+}^D \end{array}$$

commutes too.

For 4.1(4), let $h \in [\![\, \sigma \,]\!]_A^D$. Then $\bar{h} \in ([\![\, \sigma \,]\!]_A^D)^I$. For each $d \in D$, $ev_d \circ \bar{h} = \overline{h(d)} \in ([\![\, \sigma \,]\!]_A)^I$. Thus $ev_d^U(\bar{h}^U) = \overline{h(d)}^U$, which belongs to $\mathrm{Dom}\,\theta_\sigma$ by 4.1(4) for σ. Therefore by (9.vi), we get $\bar{h}^U \in \mathrm{Dom}\,\theta_{o \Rightarrow \sigma}$.

That completes the induction case for $o \Rightarrow \sigma$, and hence completes the proof of Theorem 4.1. □

9.3 Proof of Lemma 4.3

Let $x = f^U$ and $\gamma = (\gamma_1^U, \dots, \gamma_n^U)$.

(1) For each $i \in I$, put

$$h(i) = [\![\, \Gamma \rhd N \,]\!]_{\alpha_i}(f(i), g_1(i), \dots, g_n(i)).$$

Then $[\![\, \Gamma \rhd N \,]\!]^U(x, \gamma) = h^U$, so by hypothesis $h^U \in \mathrm{Dom}\,\varepsilon_j^U$. It follows that the set $I_h = \{i \in I : h(i) \in \mathrm{Dom}\,\varepsilon_{ji}\}$ belongs to U.

But for all $i \in I_h$,

$$\begin{array}{l} [\![\, \Gamma \rhd \mathsf{case}(N, M_1, M_2) \,]\!]_{\alpha_i}(f(i), g_1(i), \dots, g_n(i)) \\ = [\![\, \Gamma, v_j : \sigma_j \rhd M_j \,]\!]_{\alpha_i}(f(i), g_1(i), \dots, g_n(i), \varepsilon_j h(i)). \end{array}$$

Since $I_h \in U$, this is enough to ensure that

$$\begin{array}{l} [\![\, \Gamma \rhd \mathsf{case}(N, M_1, M_2) \,]\!]^U(f^U, g_1^U, \dots, g_n^U) \\ = [\![\, \Gamma, v_j : \sigma_j \rhd M_j \,]\!]^U(f^U, g_1^U, \dots, g_n^U, \varepsilon_j^U(h^U)), \end{array}$$

which is the desired conclusion.

(2) Define $h(i)$ as in the proof of (1), and

$$k(i) = [\![\Gamma \rhd M]\!]_{\alpha_i}(f(i), g_1(i), \dots, g_n(i)).$$

Then $[\![\Gamma \rhd M]\!]^U(x, \gamma) = k^U$ and $[\![\Gamma \rhd N]\!]^U(x, \gamma) = h^U$. Thus

$$\begin{aligned}
&[\![\Gamma \rhd M \cdot N]\!]^U(x, \gamma) \\
&= \langle [\![\Gamma \rhd M \cdot N]\!]_{\alpha_i}(f(i), g_1(i), \dots, g_n(i)) : i \in I\rangle^U \\
&= \langle k(i)(h(i)) : i \in I\rangle^U \\
&= \langle eval_i(k(i), h(i)) : i \in I\rangle^U \\
&= eval^U(k^U, h^U)
\end{aligned}$$

as required.

(3) This time define

$$h(i) = [\![\Gamma \rhd \lambda v.M]\!]_{\alpha_i}(f(i), g_1(i), \dots, g_n(i))$$

for all $i \in I$. Then

$$\begin{aligned}
&ev_d^U([\![\Gamma \rhd \lambda v.M]\!]^U(x, \gamma)) \\
&= ev_d^U(h^U) \\
&= \langle ev_{di}(h(i)) : i \in I\rangle^U \\
&= \langle [\![\Gamma, v : o \rhd M]\!]_{\alpha_i}(f(i), g_1(i), \dots, g_n(i), d) : i \in I\rangle^U \\
&= [\![\Gamma, v : o \rhd M]\!]^U(f^U, g_1^U, \dots, g_n^U, \bar{d}^U).
\end{aligned}$$

□

9.4 Proof of Theorem 5.1

The proof is by induction on the formation of the term $\Gamma \rhd M$. Part (3) is immediate from (2), since θ_σ is inverse to $e_{[\![\sigma]\!]}$ when σ is observable. Part (4) follows from (1) and (2), as θ_{St} is the inverse of the inclusion $\Pi_U A_i^+ \hookrightarrow \Pi_U A_i$. So our task is to prove (1) and (2) for each case of the axioms and rules of Figure 2, as well as the rule (τ-Tr).

The results (9.i)–(9.vii) cited below can all be found in Section 9.2.

Var Here $\Gamma \rhd M$ has the form $v : \sigma \rhd v : \sigma$. Then $[\![v : \sigma \rhd v]\!]_{\alpha_i}$ is the projection $A_i \times |\sigma|A_i \to |\sigma|A_i$, so $[\![v : \sigma \rhd v]\!]^U$ is the projection $\Pi_U A_i \times \Pi_u|\sigma|A_i \to \Pi_u|\sigma|A_i$. Also $[\![v : \sigma \rhd v]\!]_{\alpha^+}$ is the projection $\Pi_U A_i^+ \times |\sigma|\Pi_U A_i^+ \to |\sigma|\Pi_U A_i^+$. Thus for $x \in \Pi_U A_i^+$ and $\gamma \in \mathrm{Dom}\,\theta_\sigma$, $[\![v : \sigma \rhd v]\!]^U(x, \gamma) = \gamma \in \mathrm{Dom}\,\theta_\sigma$, and

$$\theta_\sigma([\![v : \sigma \rhd v]\!]^U(x, \gamma)) = \theta_\sigma(\gamma) = [\![v : \sigma \rhd v]\!]_{\alpha^+}(x, \theta_\sigma(\gamma)).$$

Con If $c \in [\![o]\!]$, then $[\![c : o]\!]_{\alpha_i} : A_i \to [\![o]\!]$ is the function with constant value c, so $[\![c : o]\!]^U(x) = \langle c : i \in I\rangle^U = e_{[\![o]\!]}(c) \in \mathrm{Dom}\,\theta_o$, and

$$\theta_o([\![c]\!]^U(x)) = c = [\![c]\!]_{\alpha^+}(x)$$

when $x \in \Pi_U A_i^+$.

St Since $[\![\mathsf{s} : \mathsf{St}]\!]_{\alpha_i} = \mathrm{id}_{A_i}$, $[\![\mathsf{s} : \mathsf{St}]\!]^U$ is the identity function on $\Pi_U A_i$. But θ_{St} is inverse to the inclusion $\Pi_U A_i^+ \hookrightarrow \Pi_U A_i$, so for $x \in \Pi_U A_i^+$, $[\![\mathsf{s}]\!]^U(x) = x \in \mathrm{Dom}\,\theta_{\mathsf{St}}$, and

$$\theta_{\mathsf{St}}([\![\mathsf{s}]\!]^U(x)) = x = [\![\mathsf{s}]\!]_{\alpha^+}(x).$$

τ-Tr Assume that 5.1(1) and 5.1(2) hold for a term $\Gamma \rhd M : \mathsf{St}$. We prove that they then hold for the term $\Gamma \rhd \mathsf{tr}(M) : \tau$.

Recall from the construction of α^+ (Theorem 4.1) that if $y \in \Pi_U A_i^+$ then $\alpha^U(y) \in \mathrm{Dom}\,\theta_\tau$ and $\theta_\tau(\alpha^U(y)) = \alpha^+(y) = \alpha^+(\theta_{\mathsf{St}}(y))$. Thus if $x \in \Pi_U A_i^+$ and $\gamma \in \mathrm{Dom}\,\theta_\Gamma$, by hypothesis 5.1(1) for $\Gamma \rhd M$,

$$y = [\![\, \Gamma \rhd M \,]\!]^U(x,\gamma) \in \mathrm{Dom}\,\theta_{\mathsf{St}} = \Pi_U A_i^+,$$

so $\alpha^U([\![\, \Gamma \rhd M \,]\!]^U(x,\gamma)) \in \mathrm{Dom}\,\theta_\tau$. But

$$\alpha^U([\![\, \Gamma \rhd M \,]\!]^U(x,\gamma)) = \langle \alpha_i \circ [\![\, \Gamma \rhd M \,]\!]_{\alpha_i} \rangle^U(x,\gamma) = [\![\, \Gamma \rhd \mathsf{tr}(M) : \tau \,]\!]^U(x,\gamma),$$

so this proves 5.1(1) for $\Gamma \rhd \mathsf{tr}(M)$. Also

$$\begin{array}{ll} \theta_\tau([\![\, \Gamma \rhd \mathsf{tr}(M) : \tau \,]\!]^U(x,\gamma)) & \\ = \theta_\tau(\alpha^U([\![\, \Gamma \rhd M \,]\!]^U(x,\gamma))) & \\ = \alpha^+(\theta_{\mathsf{St}}([\![\, \Gamma \rhd M \,]\!]^U(x,\gamma))) & \text{as above for } y \in \Pi_U A_i^+ \\ = \alpha^+([\![\, \Gamma \rhd M \,]\!]_{\alpha^+}(x,\theta_\Gamma(\gamma))) & \text{by 5.1(2) for } \Gamma \rhd M \\ = [\![\, \Gamma \rhd \mathsf{tr}(M) \,]\!]_{\alpha^+}(x,\theta_\Gamma(\gamma)), & \end{array}$$

which proves 5.1(2) for $\Gamma \rhd \mathsf{tr}(M)$.

Weak Assume the result for a term $\Gamma, \Gamma' \rhd M : \sigma$, and suppose variable v does not occur in Γ or Γ'. In line with the fact that U-liftings preserve projections, it is straightforward to show that in general

$$[\![\, \Gamma, v : \sigma', \Gamma' \rhd M : \sigma \,]\!]^U(x,\gamma,y,\gamma') = [\![\, \Gamma, \Gamma' \rhd M : \sigma \,]\!]^U(x,\gamma,\gamma').$$

It then readily follows by the induction hypothesis that 5.1(1) and 5.1(2) hold for $\Gamma, v : \sigma', \Gamma' \rhd M : \sigma$.

Proj$_j$ Assume the result for a term $\Gamma \rhd M : \sigma_1 \times \sigma_2$. If $x \in \Pi_U A_i^+$ and $\gamma \in \mathrm{Dom}\,\theta_\Gamma$, then by 5.1(1) for $\Gamma \rhd M$, we get $[\![\, \Gamma \rhd M \,]\!]^U(x,\gamma) \in \mathrm{Dom}\,\theta_{\sigma_1 \times \sigma_2}$. Then for each $j = 1, 2$, from (9.i) it follows that

$$\{\pi_j\}^U([\![\, \Gamma \rhd M \,]\!]^U(x,\gamma)) \in \mathrm{Dom}\,\theta_{\sigma_j}.$$

But using the fact that $[\![\, \Gamma \rhd \pi_j M \,]\!]_{\alpha_i} = \pi_{ji} \circ [\![\, \Gamma \rhd M \,]\!]_{\alpha_i}$ we can show that

$$[\![\, \Gamma \rhd \pi_j M \,]\!]^U(x,\gamma) = \{\pi_j\}^U([\![\, \Gamma \rhd M \,]\!]^U(x,\gamma)),$$

so this proves 5.1(1) for $\Gamma \rhd \pi_j M : \sigma_j$. Then

$$\begin{array}{ll} \theta_{\sigma_j}([\![\, \Gamma \rhd \pi_j M \,]\!]^U(x,\gamma)) & \\ = \theta_{\sigma_j}(\{\pi_j\}^U([\![\, \Gamma \rhd M \,]\!]^U(x,\gamma))) & \\ = \pi_j(\theta_{\sigma_1 \times \sigma_2}([\![\, \Gamma \rhd M \,]\!]^U(x,\gamma))) & \text{by (9.ii)} \\ = \pi_j([\![\, \Gamma \rhd M \,]\!]_{\alpha^+}(x,\theta_\Gamma(\gamma))) & \text{by 5.1(2) for } \Gamma \rhd M \\ = [\![\, \Gamma \rhd \pi_j M \,]\!]_{\alpha^+}(x,\theta_\Gamma(\gamma)), & \end{array}$$

giving 5.1(2) as well for $\Gamma \rhd \pi_j M$.

Pair Assume the result for terms $\Gamma \rhd M_j : \sigma_j$ for $j = 1$ and 2. If $x \in \Pi_U A_i^+$ and $\gamma \in \mathrm{Dom}\,\theta_\Gamma$, then for each j

$$\begin{aligned} \{\pi_j\}^U([\![\, \Gamma \rhd \langle M_1, M_2 \rangle \,]\!]^U(x,\gamma)) &= [\![\, \Gamma \rhd \pi_j \langle M_1, M_2 \rangle \,]\!]^U(x,\gamma) \\ &= [\![\, \Gamma \rhd M_j \,]\!]^U(x,\gamma), \end{aligned}$$

which belongs to $\mathrm{Dom}\,\theta_{\sigma_j}$ by 5.1(1) for $\Gamma \rhd M_j$. Therefore $[\![\, \Gamma \rhd \langle M_1, M_2 \rangle \,]\!]^U(x,\gamma)$ belongs to $\mathrm{Dom}\,\theta_{\sigma_1 \times \sigma_2}$ by (9.i), giving 5.1(1) for the term $\Gamma \rhd \langle M_1, M_2 \rangle : \sigma_1 \times \sigma_2$. Also

$$\begin{array}{ll} \pi_j[\theta_{\sigma_1 \times \sigma_2}([\![\, \Gamma \rhd \langle M_1, M_2 \rangle \,]\!]^U(x,\gamma))] & \\ = \theta_{\sigma_j}(\{\pi_j\}^U([\![\, \Gamma \rhd \langle M_1, M_2 \rangle \,]\!]^U(x,\gamma))) & \text{by (9.ii)} \\ = \theta_{\sigma_j}([\![\, \Gamma \rhd M_j \,]\!]^U(x,\gamma)) & \text{from above} \\ = [\![\, \Gamma \rhd M_j \,]\!]_{\alpha^+}(x,\theta_\Gamma(\gamma)) & \text{by 5.1(2) for } \Gamma \rhd M_j \\ = \pi_j[\,[\![\, \Gamma \rhd \langle M_1, M_2 \rangle \,]\!]_{\alpha^+}(x,\theta_\Gamma(\gamma))\,], & \end{array}$$

for $j = 1$ and 2, so

$$\theta_{\sigma_1 \times \sigma_2}(\llbracket \Gamma \rhd \langle M_1, M_2 \rangle \rrbracket^U(x,\gamma)) = \llbracket \Gamma \rhd \langle M_1, M_2 \rangle \rrbracket^U(x, \theta_\Gamma(\gamma)).$$

Hence 5.1(2) holds for $\Gamma \rhd \langle M_1, M_2 \rangle : \sigma_1 \times \sigma_2$.

$\boxed{\mathrm{In}_j}$ Assume the result for $\Gamma \rhd M : \sigma_j$. If $x \in \Pi_U A_i^+$ and $\gamma \in \mathrm{Dom}\,\theta_\Gamma$, then by 5.1(1) for $\Gamma \rhd M$, $\llbracket \Gamma \rhd M \rrbracket^U(x,\gamma) \in \mathrm{Dom}\,\theta_{\sigma_j}$, so by (9.iv),

$$\iota_j^U(\llbracket \Gamma \rhd M \rrbracket^U(x,\gamma)) \in \mathrm{Dom}\,\theta_{\sigma_1+\sigma_2}.$$

But using the fact that $\llbracket \Gamma \rhd \iota_j M \rrbracket_{\alpha_i} = \iota_{ji} \circ \llbracket \Gamma \rhd M \rrbracket_{\alpha_i}$ we can show that

$$\llbracket \Gamma \rhd \iota_j M \rrbracket^U(x,\gamma) = \iota_j^U(\llbracket \Gamma \rhd M \rrbracket^U(x,\gamma)),$$

so this proves 5.1(1) for $\Gamma \rhd \iota_j M : \sigma_1 + \sigma_2$. Then

$$\begin{aligned}
&\theta_{\sigma_1+\sigma_2}(\llbracket \Gamma \rhd \iota_j M \rrbracket^U(x,\gamma)) \\
&= \theta_{\sigma_1+\sigma_2}(\iota_j^U(\llbracket \Gamma \rhd M \rrbracket^U(x,\gamma))) \\
&= \iota_j(\theta_{\sigma_j}(\llbracket \Gamma \rhd M \rrbracket^U(x,\gamma))) && \text{by (9.v)} \\
&= \iota_j(\llbracket \Gamma \rhd M \rrbracket_{\alpha^+}(x, \theta_\Gamma(\gamma))) && \text{by 5.1(2) for } \Gamma \rhd M \\
&= \llbracket \Gamma \rhd \iota_j M \rrbracket_{\alpha^+}(x, \theta_\Gamma(\gamma)).
\end{aligned}$$

Hence 5.1(2) holds for $\Gamma \rhd \iota_j M : \sigma_1 + \sigma_2$.

$\boxed{\text{Case}}$ Here we assume the result for terms $\Gamma \rhd N : \sigma_1 + \sigma_2$ and $\Gamma, v_j : \sigma_j \rhd M_j : \sigma$ for $j = 1$ and 2, and prove it for $\Gamma \rhd \mathsf{case}(N, M_1, M_2) : \sigma$.

If $x \in \Pi_U A_i^+$ and $\gamma \in \mathrm{Dom}\,\theta_\Gamma$, by 5.1(1) for $\Gamma \rhd N$, $\llbracket \Gamma \rhd N \rrbracket^U(x,\gamma) \in \mathrm{Dom}\,\theta_{\sigma_1+\sigma_2}$, so by (9.iii) there is some j with $\llbracket \Gamma \rhd N \rrbracket^U(x,\gamma) \in \mathrm{Dom}\,\varepsilon_j^U$ and

$$\varepsilon_j^U \llbracket \Gamma \rhd N \rrbracket^U(x,\gamma) \in \mathrm{Dom}\,\theta_{\sigma_j}.$$

Then by 5.1(1) for $\Gamma, v_j : \sigma_j \rhd M_j$,

$$\llbracket \Gamma, v_j : \sigma_j \rhd M_j \rrbracket^U(x, \gamma, \varepsilon_j^U \llbracket \Gamma \rhd N \rrbracket^U(x,\gamma)) \in \mathrm{Dom}\,\theta_\sigma.$$

By Lemma 4.3(1), this means that $\llbracket \Gamma \rhd \mathsf{case}(N, M_1, M_2) \rrbracket^U(x,\gamma) \in \mathrm{Dom}\,\theta_\sigma$, as required for 5.1(1). For 5.1(2), note that by 5.1(2) for $\Gamma \rhd N$,

$$\theta_{\sigma_1+\sigma_2}(\llbracket \Gamma \rhd N \rrbracket^U(x,\gamma)) = \llbracket \Gamma \rhd N \rrbracket_{\alpha^+}(x, \theta_\Gamma(\gamma)),$$

and so by the definition of $\theta_{\sigma_1+\sigma_2}$,

$$\iota_j(\theta_{\sigma_j}(\varepsilon_j^U \llbracket \Gamma \rhd N \rrbracket^U(x,\gamma))) = \llbracket \Gamma \rhd N \rrbracket_{\alpha^+}(x, \theta_\Gamma(\gamma)). \tag{9.viii}$$

Hence $\llbracket \Gamma \rhd N \rrbracket_{\alpha^+}(x, \theta_\Gamma(\gamma)) \in \mathrm{Dom}\,\varepsilon_j$, and then

$$\begin{aligned}
&\theta_\sigma(\llbracket \Gamma \rhd \mathsf{case}(N, M_1, M_2) \rrbracket^U(x,\gamma)) \\
&= \theta_\sigma(\llbracket \Gamma, v_j : \sigma_j \rhd M_j \rrbracket^U(x, \gamma, \varepsilon_j^U \llbracket \Gamma \rhd N \rrbracket^U(x,\gamma))) && \text{Lemma 4.3(1)} \\
&= \llbracket \Gamma, v_j : \sigma_j \rhd M_j \rrbracket_{\alpha^+}(x, \theta_\Gamma(\gamma), \theta_{\sigma_j}(\varepsilon_j^U \llbracket \Gamma \rhd N \rrbracket^U(x,\gamma))) && \text{by 5.1(2) for term} \\
& && \Gamma, v_j : \sigma_j \rhd M_j \\
&= \llbracket \Gamma, v_j : \sigma_j \rhd M_j \rrbracket_{\alpha^+}(x, \theta_\Gamma(\gamma), \varepsilon_j \llbracket \Gamma \rhd N \rrbracket_{\alpha^+}(x, \theta_\Gamma(\gamma))) && \text{from (9.viii) above} \\
&= \llbracket \Gamma \rhd \mathsf{case}(N, M_1, M_2) \rrbracket_{\alpha^+}(x, \theta_\Gamma(\gamma)). && \mathsf{case}\text{ semantics}
\end{aligned}$$

This completes the proof of the result for $\Gamma \rhd \mathsf{case}(N, M_1, M_2) : \sigma$.

$\boxed{\text{App}}$ Assume the result for $\Gamma \rhd M : o \Rightarrow \sigma$ and $\Gamma \rhd N : o$. If $x \in \Pi_U A_i^+$ and $\gamma \in \mathrm{Dom}\,\theta_\Gamma$, by 5.1(1) for $\Gamma \rhd N$, $\llbracket \Gamma \rhd N \rrbracket^U(x,\gamma) \in \mathrm{Dom}\,\theta_o$. Since θ_o is inverse to the injection $\llbracket o \rrbracket \rightarrowtail \llbracket o \rrbracket^U$, there must be some $d \in \llbracket o \rrbracket$ with $\llbracket \Gamma \rhd N \rrbracket^U(x,\gamma) = \bar{d}^U$. Then by 5.1(2) for $\Gamma \rhd N$,

$$\llbracket \Gamma \rhd N \rrbracket_{\alpha^+}(x, \theta_\Gamma(\gamma)) = \theta_o(\llbracket \Gamma \rhd N \rrbracket^U(x,\gamma)) = d.$$

Now by 5.1(1) for $\Gamma \rhd M$, $[\![\Gamma \rhd M]\!]^U(x,\gamma) \in \mathrm{Dom}\,\theta_{o\Rightarrow\sigma}$, so by (9.vi),

$$ev_d^U([\![\Gamma \rhd M]\!]^U(x,\gamma)) \in \mathrm{Dom}\,\theta_\sigma.$$

But

$$\begin{aligned} & [\![\Gamma \rhd M \cdot N]\!]^U(x,\gamma) \\ = {} & eval^U([\![\Gamma \rhd M]\!]^U(x,\gamma), [\![\Gamma \rhd N]\!]^U(x,\gamma)) && \text{Lemma 4.3(2)} \\ = {} & eval^U([\![\Gamma \rhd M]\!]^U(x,\gamma), \bar{d}^U) \\ = {} & ev_d^U([\![\Gamma \rhd M]\!]^U(x,\gamma)), && \text{(4.ii)} \end{aligned}$$

so from above $[\![\Gamma \rhd M \cdot N]\!]^U(x,\gamma) \in \mathrm{Dom}\,\theta_\sigma$, which is 5.1(1) for $\Gamma \rhd M \cdot N : \sigma$. Then

$$\begin{aligned} & \theta_\sigma([\![\Gamma \rhd M \cdot N]\!]^U(x,\gamma)) \\ = {} & \theta_\sigma(ev_d^U([\![\Gamma \rhd M]\!]^U(x,\gamma))) && \text{from above} \\ = {} & \theta_{o\Rightarrow\sigma}([\![\Gamma \rhd M]\!]^U(x,\gamma))(d) && \text{(9.vii)} \\ = {} & [\![\Gamma \rhd M]\!]_{\alpha^+}(x,\theta_\Gamma(\gamma))(d) && \text{by 5.1(2) for } \Gamma \rhd M \\ = {} & [\![\Gamma \rhd M \cdot N]\!]_{\alpha^+}(x,\theta_\Gamma(\gamma)) \end{aligned}$$

as $d = [\![\Gamma \rhd N]\!]_{\alpha^+}(x,\theta_\Gamma(\gamma))$. This proves 5.1(2) for $\Gamma \rhd M \cdot N : \sigma$.

Abs Assume the result for $\Gamma, v : o \rhd M : \sigma$. If $x \in \Pi_U A_i^+$ and $\gamma \in \mathrm{Dom}\,\theta_\Gamma$, then for any $d \in [\![o]\!]$, by Lemma 4.3(3)

$$ev_d^U([\![\Gamma \rhd \lambda v.M : o \Rightarrow \sigma]\!]^U(x,\gamma)) = [\![\Gamma, v : o \rhd M]\!]^U(x,\gamma,\bar{d}^U),$$

which belongs to $\mathrm{Dom}\,\theta_\sigma$ by 5.1(1) for $\Gamma, v : o \rhd M$, since $\bar{d}^U \in \mathrm{Dom}\,\theta_o$. By (9.vi) this means that $[\![\Gamma \rhd \lambda v.M]\!]^U(x,\gamma) \in \mathrm{Dom}\,\theta_{o\Rightarrow\sigma}$, so 5.1(1) holds for $\Gamma \rhd \lambda v.M$. Then

$$\begin{aligned} & \theta_{o\Rightarrow\sigma}([\![\Gamma \rhd \lambda v.M]\!]^U(x,\gamma))(d) \\ = {} & \theta_\sigma(ev_d^U([\![\Gamma \rhd \lambda v.M]\!]^U(x,\gamma))) && \text{by (9.vii)} \\ = {} & \theta_\sigma([\![\Gamma, v : o \rhd M]\!]^U(x,\gamma,\bar{d}^U)) && \text{Lemma 4.3(3)} \\ = {} & [\![\Gamma, v : o \rhd M]\!]_{\alpha^+}(x,\theta_\Gamma(\gamma),\theta_o(\bar{d}^U)) && \text{by 5.2(2) for } \Gamma, v : o \rhd M \\ = {} & [\![\Gamma, v : o \rhd M]\!]_{\alpha^+}(x,\theta_\Gamma(\gamma),d) \\ = {} & [\![\Gamma \rhd \lambda v.M]\!]_{\alpha^+}(x,\theta_\Gamma(\gamma))(d) \end{aligned}$$

for any $d \in [\![o]\!]$. Hence

$$\theta_{o\Rightarrow\sigma}([\![\Gamma \rhd \lambda v.M]\!]^U(x,\gamma)) = [\![\Gamma \rhd \lambda v.M]\!]_{\alpha^+}(x,\theta_\Gamma(\gamma)),$$

proving 5.1(2) for $\Gamma \rhd \lambda v.M$.

This completes the proof of Theorem 5.1. □

References

[1] Peter Aczel and Nax Mendler. A final coalgebra theorem. In D. H. Pitt et al., editors, *Category Theory and Computer Science. Proceedings 1989*, volume 389 of *Lecture Notes in Computer Science*, pages 357–365. Springer-Verlag, 1989.

[2] Jiří Adámek and Hans-E. Porst. From varieties of algebras to covarieties of coalgebras. *Electronic Notes in Theoretical Computer Science*, 44, Issue 1, 2001. `http://www.elsevier.nl/locate/entcs`.

[3] Steve Awodey and Jesse Hughes. The coalgebraic dual of Birkhoff's variety theorem. Technical Report CMU-PHIL-109, Department of Philosophy, Carnegie Mellon University, 2000.

[4] Steve Awodey and Jesse Hughes. Modal operators and the formal dual of Birkhoff's completeness theorem. *Mathematical Structures in Computer Science*, 13:233–258, 2003.

[5] B. Banaschewski and H. Herrlich. Subcategories defined by implications. *Houston Journal of Mathematics*, 2(2):149–171, 1976.

[6] J. L. Bell and A. B. Slomson. *Models and Ultraproducts.* North-Holland, Amsterdam, 1969.

[7] Garrett Birkhoff. On the structure of abstract algebras. *Proceedings of the Cambridge Philosophical Society*, 31:433–454, 1935.

[8] S. Burris and H. P. Sankappanavar. *A Course in Universal Algebra.* Springer-Verlag, 1981.

[9] C. C. Chang and H. J. Keisler. *Model Theory.* North-Holland, Amsterdam, 1973.

[10] Robert Goldblatt. Equational logic of polynomial coalgebras. In *Advances in Modal Logic*, volume 4, World Scientific (in press). Manuscript available at `http://www.mcs.vuw.ac.nz/~rob`

[11] Robert Goldblatt. What is the coalgebraic analogue of Birkhoff's variety theorem? *Theoretical Computer Science*, 266:853–886, 2001.

[12] G. Grätzer. *Universal Algebra.* D. Van Nostrand, 1968.

[13] H. Peter Gumm. Elements of the general theory of coalgebras. LUATCS'99, Rand Africaans University, Johannesburg, South Africa, 60 pp. Electronically available at `http://www.Mathematik.uni-marburg.de/~gumm/Papers/publ.html`, 1999.

[14] H. Peter Gumm. Birkhoff's variety theorem for coalgebras. *Contributions to General Algebra*, 13:159–173, 2000.

[15] H. Peter Gumm and Tobias Schröder. Covarieties and complete covarieties. *Electronic Notes in Theoretical Computer Science*, 11, 1998. `http://www.elsevier.nl/locate/entcs`.

[16] William S. Hatcher. Quasiprimitive subcategories. *Mathematische Annalen*, 190:93–96, 1970.

[17] Matthew Hennessy and Robin Milner. Algebraic laws for nondeterminism and concurrency. *Journal of the Association for Computing Machinery*, 32:137–161, 1985.

[18] Claudio Hermida. *Fibrations, Logical Predicates and Indeterminates.* PhD thesis, University of Edinburgh, 1993. Techn. rep. LFCS-93-277. Also available as Aarhus Univ. DAIMI Techn. rep. PB-462.

[19] Claudio Hermida and Bart Jacobs. Structural induction and coinduction in a fibrational setting. *Information and Computation*, 145:107–152, 1998.

[20] H. Herrlich and C. M. Ringel. Identities in categories. *Canadian Mathematical Bulletin*, 15(2):297–299, 1972.

[21] Jesse Hughes. *A Study of Categories of Algebras and Coalgebras.* PhD thesis, Carnegie Mellon University, 2001.

[22] Bart Jacobs. Mongruences and cofree coalgebras. In V. S. Alagar and M. Nivat, editors, *Algebraic Methods and Software Technology*, volume 936 of *Lecture Notes in Computer Science*, pages 245–260. Springer, 1995.

[23] Bart Jacobs. Objects and classes, coalgebraically. In B. Freitag, C. B. Jones, C. Lengauer, and H.-J. Schek, editors, *Object-Orientation with Parallelism and Persistence*, pages 83–103. Kluwer Academic Publishers, 1996.

[24] Bart Jacobs. Towards a duality result in coalgebraic modal logic. *Electronic Notes in Theoretical Computer Science*, 33, 2000. `http://www.elsevier.nl/locate/entcs`.

[25] Bart Jacobs. Exercises in coalgebraic specification. In R. Backhouse, R. Crole, and J. Gibbons, editors, *Algebraic and Coalgebraic Methods in the Mathematics of Program Construction*, volume 2297 of *Lecture Notes in Computer Science*, pages 237–280. Springer, 2002.

[26] Bart Jacobs and Jan Rutten. A tutorial on (co)algebras and (co)induction. *Bulletin of the European Association for Theoretical Computer Science*, 62:222–259, 1997.

[27] Alexander Kurz. Coalgebras and modal logic. In *Proceedings of Advances in Modal Logic '98, Uppsala*, pages 222–230, 1998.

[28] Alexander Kurz. *Logic for Coalgebras and Applications to Computer Science*. PhD thesis, Ludwig-Maximilians-Universität München, 2000.

[29] Alexander Kurz. A covariety theorem for modal logic. In Michael Zakharyaschev, Krister Segerberg, Maarten de Rijke, and Heinrich Wansang, editors, *Advances in Modal Logic, volume 2*, pages 367–380. CSLI Publications, 2001.

[30] David Park. Concurrency and automata on infinite sequences. In P. Deussen, editor, *Theoretical Computer Science*, volume 104 of *Lecture Notes in Computer Science*, pages 167–183. Springer-Verlag, 1981.

[31] Andrew M. Pitts. Categorical logic. In S. Abramsky, D. M. Gabbay, and T. S. E. Maibaum, editors, *Handbook of Logic in Computer Science, Volume 5: Algebraic and Logical Structures*, chapter 2. Oxford University Press, 2000.

[32] Horst Reichel. An approach to object semantics based on terminal co-algebras. *Mathematical Structures in Computer Science*, 5:129–152, 1995.

[33] Grigore Roşu. A Birkhoff-like axiomatisability result for hidden algebra and coalgebra. *Electronic Notes in Theoretical Computer Science*, 11, 1998. `http://www.elsevier.nl/locate/entcs`.

[34] J.J.M.M. Rutten. A calculus of transition systems (towards universal coalgebra). In Alban Ponse, Maarten de Rijke, and Yde Venema, editors, *Modal Logic and Process Algebra*, CSLI Lecture Notes No. 53, pages 231–256. CSLI Publications, Stanford, California, 1995.

[35] J.J.M.M. Rutten. Universal coalgebra: a theory of systems. Technical Report CS-R9652, Centrum voor Wiskunde en Informatica (CWI), Amsterdam, 1996.

[36] J.J.M.M. Rutten. Universal coalgebra: a theory of systems. *Theoretical Computer Science*, 249(1):3–80, 2000.