Available online at www.sciencedirect.com
SCIENCE DIRECT®
Annals of Pure and Applied Logic 138 (2006) 77–93
ANNALS OF PURE AND APPLIED LOGIC
www.elsevier.com/locate/apal

Final coalgebras and the Hennessy–Milner property

Robert Goldblatt

Centre for Logic, Language and Computation, Victoria University of Wellington, New Zealand

Available online 11 July 2005

Abstract

The existence of a final coalgebra is equivalent to the existence of a formal logic with a set (small class) of formulas that has the Hennessy–Milner property of distinguishing coalgebraic states up to bisimilarity. This applies to coalgebras of any functor on the category of sets for which the bisimilarity relation is transitive. There are cases of functors that do have logics with the Hennessy–Milner property, but the only such logics have a proper class of formulas.

The main theorem gives a representation of states of the final coalgebra as certain satisfiable sets of formulas. The key technical fact used is that any function between coalgebras that is truth-preserving and has a simple codomain must be a coalgebraic morphism.

Keywords: Coalgebra; Final object; Bisimulation; Bisimilarity; Congruence; Coinductive; Abstract logic

1. Introduction and overview

Coalgebras of functors $T : \mathbf{Set} \to \mathbf{Set}$ on the category of sets have proven useful in modelling notions associated with data structures, transition systems, and process algebras [19,9,22,24,11]. Of particular importance is the notion of a *final* (or terminal) coalgebra, which is a coalgebra γ such that for each coalgebra α there is a unique coalgebraic morphism from α to γ. If such a final coalgebra exists, its members can be thought of as representing all possible "behaviours" of processes, because members x and y of coalgebras α and β (respectively) are typically "behaviourally indistinguishable" precisely when they are identified by the unique morphisms from α and β to γ (see 3.6 below).

E-mail address: Rob.Goldblatt@vuw.ac.nz.

0168-0072/$ - see front matter
doi:10.1016/j.apal.2005.06.006

This paper addresses the following question:

> when can logic be used to construct a final object in the category of coalgebras of a functor $T : \mathbf{Set} \to \mathbf{Set}$?

We will demonstrate an abstract connection between this question and the issue of whether there exists a logical system, comprising a set of formulas with a semantics, that distinguishes states of T-coalgebras up to *bisimilarity*, meaning that two states satisfy the same formulas of the language precisely when they are bisimilar.

Specifically, we prove that a final T-coalgebra exists if there exists a logic with this property of having a set of formulas that differentiates coalgebraic elements up to bisimilarity. Conversely, the existence of a final coalgebra implies the existence of a logical system with this property, provided that the bisimilarity relation between coalgebraic states is transitive. Such transitivity certainly holds in any process-algebraic context for which bisimilarity means *observational (or behavioural) indistinguishability*.

Hennessy and Milner [7,8] introduced into the study of computational processes the seminal idea of associating with a given type of state-transition system a logic that has this fundamental property that two states are observationally indistinguishable (or bisimilar) precisely when they are *logically* indistinguishable in the sense of satisfying the same formulas of the language. They showed that for finitely branching systems this objective is realized by a finitary modal language, while for systems in general it can be realized by using an infinitary language that allows conjunctions and disjunctions of any set of formulas (see Example 4.2). Note that for such an infinitary language the collection of all formulas may constitute a proper class, rather than a set. This point is pivotal to our discussion, as we will see that the existence of a final coalgebra depends on the existence of a suitable logic whose formulas constitute a *set*.

The Hennessy–Milner program has been applied to many species of process algebra (see [17,4]), and was extended to certain kinds of coalgebra over **Set** once it was recognized that such coalgebras can be used effectively to model data types and transition systems [19,9,22,24,11]. In that setting, final coalgebras are important both for providing operational semantics for certain languages [25] and for providing definitions of entities, and proofs of their properties, by the principle of coinduction [18,11,24].

Finitary modal languages have been devised that are expressive enough to distinguish coalgebraic elements up to bisimilarity when the functor T is *polynomial* [20,13]. A polynomial functor is one that is constructed from constant-valued functors and the identity functor by forming binary products and coproducts, and exponentials with constant exponent. In the polynomial case T-coalgebras can be thought of as modelling transition systems that are deterministic, with the constant-valued functors corresponding to sets of "outputs". The well known *canonical model* construction from modal logic [15,26] is used in [20,13] to build certain polynomial coalgebras. The essence of this method is to define a model whose elements are special sets of formulas with properties determined by the logic, and to show that a formula φ is satisfied in this model by an element x iff $\varphi \in x$. The technique was used in [20,13] to construct final polynomial coalgebras under the strong restriction that any constant functor involved in the formation of T has a *finite* output-set as its constant value.

To model non-determinism, the class of polynomial functors must be extended to allow powerset formation. Modal languages are also available for this extension [21,10], but the full use of the powerset functor $\mathcal{P}$ prevents there being any final coalgebra at all. On the other hand, we can model *finitely branching* non-determinism by using the finitary powerset functor $\mathcal{P}_\omega$, where $\mathcal{P}_\omega A$ is the set of all finite subsets of A, and for this there is a kind of modal canonical model construction of final coalgebras [10, Theorem 5.8]. Again this only works under the restriction to finite constant sets, and it appears that for the approach of [10] this finiteness restriction is also needed to establish the Hennessy–Milner principle that logically equivalent states are bisimilar.

This raises the question of whether a version of the canonical model method can produce final T-coalgebras when T involves *infinite* constant sets. The main result of this paper (Theorem 4.3) gives a very general answer by showing that a straightforward logical construction of final coalgebras is always possible whenever there exists a formal language associated with T with a set of formulas having the Hennessy–Milner property. Moreover, the construction connects with the canonical model idea on an abstract level.

Thus the Hennessy–Milner property has a powerful impact on the structure of a category of coalgebras. As part of the proof that our construction gives a final coalgebra, we use this property (in Theorem 4.1) to show that if a function between the state sets of two coalgebras leaves invariant the truth-value of formulas, and the codomain of that function is a *simple* coalgebra, then the function must be a coalgebraic *morphism*. Simple coalgebras are those that have no proper epimorphic images. They have the *coinductive* property that distinct states are never bisimilar. As part of our analysis we study the relationship between coalgebras that are simple, which equivalently means "no non-trivial congruences", and those that are coinductive, which means "no non-trivial bisimulations". It turns out that if the bisimilarity relation for T-coalgebras is transitive, the coinductive coalgebras coincide with the simple ones (Theorem 3.3). The Hennessy–Milner property itself guarantees this transitivity.

Section 2 of the paper reviews the basic theory of coalgebras and bisimulations that will be used. Section 3 discusses the construction of simple coalgebras by the theory of congruences, and explains their relationship to coinductive coalgebras. Section 4 formulates the abstract notion of a logic for coalgebras and gives the main result characterizing the existence of final coalgebras. Section 5 gives a formulation in the case that bisimilarity is not transitive. The paper concludes with some further questions.

2. Coalgebras and their bisimulations

Fix a covariant endofunctor $T : \mathbf{Set} \to \mathbf{Set}$ on the category of sets and functions. A T-*coalgebra* is a pair (A, α) comprising a set A and a function α of the form $A \to TA$. A will be called the set of *states* and α the *transition structure* of the coalgebra. We often identify a coalgebra with its transition structure α (from which the state set A can be determined as the domain of α). A *pointed coalgebra* is a pair (α, x) with x being a state of coalgebra α.

A T-*morphism* from T-coalgebra (A, α) to T-coalgebra (B, β) is a function $f : A \to B$ 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:

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ {\scriptstyle\alpha}\downarrow & & \downarrow{\scriptstyle\beta} \\ TA & \xrightarrow{Tf} & TB \end{array}$$

The class of T-coalgebras with their T-morphisms form a category T-**Coalg** under functional composition of morphisms. A *final* object in this category is a T-coalgebra β such that for each T-coalgebra α there is *exactly one* morphism from α to β. Such unique morphisms to a final coalgebra provide coinductive definitions of many operations of importance in the study of data structures [19,11].

If (A, α) is a final coalgebra, then $\alpha : A \to TA$ is an isomorphism in **Set**, i.e. a bijection. This is Lambek's Lemma [14].

Example 2.1. A non-deterministic state-transition system (A, I, τ) has a set A of states, a set I of inputs and a relation $\tau \subseteq A \times I \times A$. Write $x \overset{i}{\mapsto} y$ when $(x, i, y) \in \tau$, signifying that there is a possible transition from x to y on input i. Putting $\alpha(x) = \{(i, y) : x \overset{i}{\mapsto} y\}$ makes the system into a coalgebra $\alpha : A \to \mathcal{P}(I \times A)$ for the functor $\mathcal{P}(I \times -)$, where $\mathcal{P}$ is the covariant powerset functor on **Set**. We may alternatively view it as a $(\mathcal{P}-)^I$-coalgebra, by taking $\alpha(x) \in (\mathcal{P}A)^I$ to be the function $\alpha(x)(i) = \{y : x \overset{i}{\mapsto} y\}$. A coalgebraic morphism is characterized as a function satisfying

$$f(x) \overset{i}{\mapsto} z \quad \text{iff} \quad \exists y (x \overset{i}{\mapsto} y \text{ and } f(y) = z).$$

Lambeck's Lemma tells us that there is no final $(\mathcal{P}-)^I$-coalgebra, since there can be no bijection of the form $A \to (\mathcal{P}A)^I$, even when I has one element, by Cantor's theorem. On the other hand a final coalgebra will exist for systems that model *finitely branching* non-determinism, which means that the set $\{y : x \overset{i}{\mapsto} y\}$ is finite for all pairs (x, i). Such a system may be viewed as a $(\mathcal{P}_\omega-)^I$-coalgebra, where $\mathcal{P}_\omega A$ is the set of all finite subsets of A. The category $(\mathcal{P}_\omega-)^I$-**Coalg** does have a final object [3,24]. $\square$

The first comprehensive study of T-**Coalg** was made by Rutten [23,24], showing that many interesting results follow if T weakly preserves certain limits,[1] an assumption that is satisfied by most functors used in practice to describe data structures. The theory without this assumption has been explored by Gumm and Schröder [5,6] and we will call on many results from these references. The situation is subtle, and needs careful attention, particularly in relation to the distinction between "congruences" and "bisimulations", as we will see.

In the category T-**Coalg**, a morphism is *epi* (right-cancellative) iff it is surjective as a set function. Any injective morphism is *mono* (left-cancellative), but the converse need not be true. In fact a morphism is injective precisely when it is an equalizer [6, 3.4]. To have every mono being injective requires some condition on T, such as that it weakly preserve the pullback of any morphism with itself [6, 5.5].

[1] A *weak* version of a type of limit is an entity that satisfies the existence part of the definition of that type of limit, but not necessarily the uniqueness part. A functor weakly preserves a type of limit if it maps any instance of that limit to a weak version of it.

A morphism is *iso* if it has an inverse in T**-Coalg**, in which case it is bijective. But the inverse of a *bijective* morphism is also a morphism [24, 2.3], so an isomorphism is the same thing as a bijective morphism.

A set B of states of coalgebra α is *closed in* α if there exists a transition $\beta : B \to TB$ for which the inclusion function from B to A is a T-morphism from β to α. If such a β exists it is unique, and in that case (B, β), or just B, is a *subcoalgebra* of α. For any morphism f with domain (A, α), the image set $Im\, f = f(A) = \{f(x) : x \in A\}$ is a subcoalgebra of the codomain of f.

A coproduct $\Sigma_{i \in I}\alpha_i$ exists in T**-Coalg** for any set of coalgebras (A_i, α_i). Its state set is the disjoint union $\Sigma_{i \in I} A_i$ of the state sets of the coalgebras. For each $j \in I$ there is an injective *insertion* function $\iota_j : A_j \to \Sigma_I A_i$, with each member of $\Sigma_I A_i$ being equal to $\iota_j(x)$ for a unique $j \in I$ and a unique $x \in A_j$. The transition structure on the coproduct acts as α_j on the image of ι_j: more precisely it acts as $\iota_j(x) \mapsto T(\iota_j)(\alpha_j(x))$. The insertion ι_j is a morphism from α_j to $\Sigma_I \alpha_i$ making α_j isomorphic to the subcoalgebra $Im\, \iota_j$ of $\Sigma_I \alpha_i$.

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 T-morphisms from ρ to α and β, i.e. the following diagram commutes:

$$\begin{tikzcd} A \arrow[d, "\alpha"'] & R \arrow[l, "\pi_1"'] \arrow[r, "\pi_2"] \arrow[d, "\rho"] & B \arrow[d, "\beta"] \\ TA & TR \arrow[l, "T\pi_1"'] \arrow[r, "T\pi_2"] & TB \end{tikzcd}$$

This definition of bisimulation was introduced in [2]. It gives a categorical formulation of a notion that has various manifestations in different kinds of state-transition system.

The union of any collection of bisimulations from α to β is a bisimulation. Hence there is a *largest* such bisimulation (the union of all of them), which is a relation known as *bisimilarity*. This will be denoted $\sim_{\alpha\beta}$, or $\sim_\alpha$ when $\alpha = \beta$, and may be written without any subscript if the coalgebras involved are understood. We may also write $(\alpha, x) \sim (\beta, y)$ when $x \sim_{\alpha\beta} y$, and view this as a relation between pointed coalgebras.

In process algebra, the existence of a bisimulation relating a pair of states x, y captures the idea that x and y are *observationally indistinguishable*. So the observational indistinguishability relation itself is identified with bisimilarity.

It is immediate from the definition of bisimulation that if $x \sim_{\alpha\beta} y$, then there exists a coalgebra ρ with morphisms $g : \rho \to \alpha$ and $h : \rho \to \beta$, and a state z of ρ such that $x = g(z)$ and $y = h(z)$. But the converse is also true, since if g and h are morphisms with the same domain, then the set of pairs $(g(z), h(z))$ for all states z of this domain is a bisimulation [24, 5.3]. It follows that *bisimilarity is preserved by morphisms*, in the sense that for any T-morphisms $f : \alpha \to \gamma$ and $g : \beta \to \delta$, if $x \sim_{\alpha\beta} y$ then $f(x) \sim_{\gamma\delta} g(y)$. In fact it can be shown, more strongly, that if R is a bisimulation from α to β, then $\{(f(x), g(y)) : xRy\}$ is a bisimulation from γ to δ.

A function $f : A \to B$ is a morphism from α to β iff its graph $\{(x, f(x)) : x \in A\}$ is a bisimulation from α to β [24, 2.5]: a morphism is essentially a functional bisimulation. Thus $(\alpha, x) \sim (\beta, f(x))$ whenever f is a morphism.

When $A \subseteq B$, α is a subcoalgebra of β iff the identity relation Δ_A on A is a bisimulation from α to β. Δ_A is itself always a bisimulation from α to α, so the bisimilarity relation $\sim_\alpha$ on α is reflexive: $\Delta_A \subseteq \sim_\alpha$. It is also symmetric, as the inverse of any bisimulation is a bisimulation. The latter property shows more generally that bisimilarity is symmetric as a global relation between pointed coalgebras, i.e. $x \sim_{\alpha\beta} y$ implies $y \sim_{\beta\alpha} x$. It need not however be transitive. So, we will say that T *has transitive bisimilarity* if

$$x \sim_{\alpha\beta} y \sim_{\beta\gamma} z \quad \text{implies} \quad x \sim_{\alpha\gamma} z$$

for all pointed T-coalgebras (α, x), (β, y), (γ, z). When this holds, $\sim_\alpha$ is an equivalence relation on each coalgebra α. Transitivity of bisimilarity would follow readily if the relational composition of two bisimulations was also a bisimulation. But closure of bisimulations under composition itself holds iff T weakly preserves pullbacks [6, 5.1]. In Theorem 3.7 below it is shown that weak preservation of a restricted class of pullbacks suffices for bisimilarity to be transitive.

If α is a final T-coalgebra, then the bisimilarity relation on α is just the identity relation (as will also be explained below). This is the basis of the proof principle of *coinduction*, which states that to prove two states of a final coalgebra equal it suffices to show that there is a bisimulation that relates them [16,18,11]. It also suggests that a final coalgebra might be constructible by some process of identification of bisimilar states. A potential obstacle here is that $\sim$ need not be an equivalence relation in general. But even when it is an equivalence there may not be any final coalgebra, as we show for the case $T = \mathcal{P}$ in Example 3.8. So some additional property will be needed.

3. Simple versus coinductive

We are going to build a final coalgebra by a quotient construction, so we first review the coalgebraic approach to these.

Let θ be an equivalence relation on a set A, with equivalence classes $x^\theta = \{y \in A : x\theta y\}$, quotient set $A/\theta = \{x^\theta : x \in A\}$, and quotient map $f_\theta : A \to A/\theta$ having $f_\theta(x) = x^\theta$. In classical universal algebra, if A is the underlying set of some algebra, then θ is called a *congruence* if the algebraic structure can be transfered to A/θ to make the map f_θ a homomorphism. Now any function f has the *kernel* equivalence relation

$$Ker\, f = \{(x, y) : f(x) = f(y)\}$$

on its domain, and when f is a homomorphism of algebras this is a congruence. Since $Ker\, f_\theta = \theta$, the congruences are thus just the kernels of homomorphisms, and any homomorphic image of an algebra is shown to be isomorphic to the quotient of that algebra by its kernel congruence. An algebra is *simple* if it has no non-trivial congruence relations, i.e. no congruences other than the identity relation Δ_A and the universal relation $A \times A$. Equivalently, this means that the algebra has no proper homomorphic images: every epimorphism with that algebra as domain either identifies all elements or is an isomorphism.

Suppose instead that A is the state set of a T-coalgebra α. What does it take to make the quotient set A/θ into a T-coalgebra? The answer, given in [2], is that

$$\theta \subseteq Ker(T(f_\theta) \circ \alpha).$$

This condition is satisfied by any bisimulation on α [2, 6.1]. For an equivalence relation θ, the following are equivalent [5, 4.12]:

(i) $\theta \subseteq Ker(T(f_\theta) \circ \alpha)$;
(ii) there is a (unique) transition structure $\alpha_\theta : A/\theta \to T(A/\theta)$ for which the quotient map $f_\theta : A \to A/\theta$ is a T-morphism from α to the coalgebra $\alpha/\theta = (A/\theta, \alpha_\theta)$, viz. $\alpha_\theta(x^\theta) = (T(f_\theta) \circ \alpha)(x)$;
(iii) θ is the kernel of some morphism in T-**Coalg** with domain α.

An equivalence relation satisfying these conditions is called a *congruence* on the coalgebra α. Thus any equivalence relation that is a bisimulation must be a congruence. But the converse, that every congruence is a bisimulation, is true iff T weakly preserves kernels [2,6], a condition that is defined just before Theorem 3.7 below.

Each bisimulation R on α has a smallest extension to a congruence θ on α, as shown in [2] and [5, 5.15]. If θ is the equivalence relation generated by R, then there is a T-transition α_θ on A/θ such that the quotient map $f_\theta : A \to A/\theta$ is a T-morphism from α to α_θ. The existence of α_θ depends on the fact that f_θ coequalizes the pair of projections $\pi_1, \pi_2 : R \to A$ in **Set** when R generates θ. Then θ, being the kernel of f_θ, is a congruence.

The set of all congruences on a coalgebra α is a complete lattice under the partial ordering $\subseteq$ of set inclusion. In particular there is a smallest congruence, namely Δ_A, and a largest congruence ∇_α. But unlike the universal algebra case, ∇_α may be smaller than the universal relation $A \times A$.

A coalgebra α will be called *simple* if its largest (and hence only) congruence is Δ_A, i.e. if $\nabla_\alpha = \Delta_A$.

Theorem 3.1. *The following are equivalent.*

(1) *α is a simple coalgebra.*
(2) *Every morphism with domain α is injective.*
(3) *Every epimorphism with domain α is an isomorphism.*

Proof. (1) implies (2): for any morphism $f : \alpha \to \beta$, the kernel $Ker\, f$ is a congruence on α, so if α is simple then $Ker\, f = \Delta_A$, which implies that f is injective.

(2) implies (3): If $f : \alpha \to \beta$ is a surjective morphism, then by (2) f is bijective. But a bijective morphism is an isomorphism [24, 2.3].

(3) implies (1): the quotient map $f_{\nabla_\alpha} : \alpha \to \alpha/\nabla_\alpha$ is an epimorphism, hence by (3) is injective, so $Ker\, f_{\nabla_\alpha} = \Delta_A$. But $Ker\, f_{\nabla_\alpha} = \nabla_\alpha$, so α is simple.

Corollary 3.2. (1) *For any coalgebra α, α/∇_α is simple.*
(2) *Every final coalgebra is simple.*

Proof. (1) Let g be any morphism with domain α/∇_α. Abbreviate ∇_α to ∇. Composing with the quotient morphism $f_\nabla : \alpha \to \alpha/\nabla$, we get that $Ker(g \circ f_\nabla)$ is a congruence on α that includes ∇, and hence is equal to ∇. Thus $g(x^\nabla) = g(y^\nabla)$ implies $x \nabla y$ and hence $x^\nabla = y^\nabla$. So g is injective, and simplicity of α follows from 3.1(2).
(2) Let α be a final T-coalgebra. For any morphism $f : \alpha \to \beta$, by finality there is a morphism $g : \beta \to \alpha$, so by finality again $g \circ f$ must be the identity function on α. Thus f has a left inverse, implying that it is injective. Hence α is simple by 3.1(2). $\square$

A coalgebra α will be called *coinductive*[2] if its largest *bisimulation* is the identity relation Δ_A, i.e. if $\sim_\alpha = \Delta_A$. In [5, 6.13] it is shown that for any coalgebra α the following are equivalent:

(iv) α is coinductive;
(v) every morphism with domain α is mono;
(vi) for any T-coalgebra β there is *at most* one morphism $\beta \to \alpha$.

Now $\sim_\alpha$ can be extended to a congruence, since it is a bisimulation, and so we always have $\Delta_A \subseteq \sim_\alpha \subseteq \nabla_\alpha$. It follows that *every simple coalgebra is coinductive*. Hence by Corollary 3.2(1), α/∇_α is always coinductive, and so by the equivalence of (iv) and (vi),

(vii) for any T-coalgebras α, β there is *at most* one morphism $\beta \to \alpha/\nabla_\alpha$.

If the largest congruence ∇_α is a bisimulation, then it is included in the largest bisimulation $\sim_\alpha$, and so altogether $\nabla_\alpha = \sim_\alpha$. If that always holds, then all coinductive coalgebras are simple. It is known that weak preservation of kernels by T will ensure this, since in that case every congruence is a bisimulation [6, 5.3]. Here is an alternative sufficient condition that is relevant to our main result.

Theorem 3.3. *If T has transitive bisimilarity, then $\sim_\alpha = \nabla_\alpha$ for all T-coalgebras α, and so every coinductive coalgebra is simple.*

Proof. If $x \nabla y$ in α, then $f_\nabla(x) = f_\nabla(y)$ in α/∇, so as f_∇ is a morphism and bisimilarity is symmetric, $x \sim f_\nabla(x) = f_\nabla(y) \sim y$. Transitivity of bisimilarity then gives $x \sim y$. This shows $\nabla_\alpha \subseteq \sim_\alpha$.

To further study the condition $\sim_\alpha = \nabla_\alpha$, we will say that *$T$-subcoalgebras preserve bisimilarity* if, whenever α is a subcoalgebra of T-coalgebra β, then $x \sim_\beta y$ implies $x \sim_\alpha y$ for all $x, y \in A$. (Since $x \sim_\alpha y$ implies $x \sim_\beta y$, because bisimilarity is preserved by the inclusion morphism $i : \alpha \to \beta$, this amounts to saying that $\sim_\alpha$ is just the restriction of $\sim_\beta$ to $A \times A$.)

We can also say that *T-subcoalgebras preserve* ∇ if, similarly, $x \nabla_\beta y$ implies $x \nabla_\alpha y$ for all $x, y \in A$ when α is a subcoalgebra of β. *But this is always true for any T*: the composition of the inclusion morphism $i : \alpha \to \beta$ with the quotient morphism $f_{\nabla_\beta} : \beta \to \beta/\nabla_\beta$ is a morphism $\alpha \to \beta/\nabla_\beta$ whose kernel congruence is a subset of ∇_α. But if $x, y \in A$ and $x \nabla_\beta y$, then (x, y) belongs to this kernel, and so $x \nabla_\alpha y$.

Theorem 3.4. *For any functor T, the following are equivalent.*

(1) $\sim_\alpha = \nabla_\alpha$ *for all T-coalgebras α.*
(2) *T-subcoalgebras preserve bisimilarity, and $\sim_\alpha$ is transitive for all α.*

Proof. Assume (1). Then (2) is immediate because subcoalgebras preserve ∇, as just noted, and ∇_α is transitive.

[2] This "coinductive" concept is called "simple" in [24,5], while our "simple" is called "strongly simple" in [5] and "s-extensional" in [2]. The use of "simple" in this paper is intended to parallel its standard use in algebra.

Conversely, assume (2) and let $x \nabla_\alpha y$. Let β be the coproduct $\alpha + (\alpha/\nabla_\alpha) + \alpha$, with $\iota : \alpha \to \beta$ being the insertion morphism into the left summand of β. In [6, 5.8] it is shown that transitivity of $\sim_\beta$ implies that $\iota(x) \sim_\beta \iota(y)$. Now $\iota(A)$ is a subcoalgebra of β, so we get $\iota(x) \sim_{\iota(A)} \iota(y)$ as T-subcoalgebras preserve bisimilarity. Since ι is an isomorphism between α and $\iota(A)$, it follows readily that $\{(z, w) : \iota(z) \sim_{\iota(A)} \iota(w)\}$ is a bisimulation on α containing (x, y), so $x \sim_\alpha y$. □

Example 3.5. There is an example in [2, p. 363] of a functor T having a two-state coalgebra $(\{0, 1\}, \alpha)$ that has $\sim_\alpha = \Delta_{\{0,1\}}$, so α is coinductive, but $\nabla_\alpha = \{0, 1\} \times \{0, 1\}$, so α is not simple. It follows from Theorem 3.3 that this T does not have transitive bisimilarity. On the other hand it does have a (one-state) final coalgebra. T acts on sets by $TA = \{(x, y, z) \in A^3 : |\{x, y, z\}| < 3\}$, and on functions by $Tf(x, y, z) = (fx, fy, fz)$. The unique function $\{0\} \to T\{0\} = \{(0, 0, 0)\}$ is a final T-coalgebra. The two-state example has $\alpha(0) = (0, 0, 1)$ and $\alpha(1) = (0, 1, 1)$. There is no bisimulation on α relating 0 to 1. □

If a functor T does have a final coalgebra γ, then each bisimilarity relation $\sim_\alpha$ is characterized as the kernel of the unique morphism $\alpha \to \gamma$, provided that T obeys some restriction, such as weak preservation of pullbacks [24, 9.3]. Here is a refined analysis of the situation that will also figure in our main result.

Theorem 3.6. *Suppose that γ is a final T-coalgebra, with a unique morphism $f_\alpha : \alpha \to \gamma$ for each T-coalgebra α. Then for any pointed T-coalgebras (α, x) and (β, y),*

(1) *$x \sim_{\alpha\beta} y$ implies $f_\alpha(x) = f_\beta(y)$; and*
(2) *if T has transitive bisimilarity, then $f_\alpha(x) = f_\beta(y)$ implies $x \sim_{\alpha\beta} y$.*

Proof. (1) if $x \sim_{\alpha\beta} y$, then (x, y) belongs to some bisimulation R from α to β. Hence there exists a coalgebra ρ on R such that the projections give morphisms $\pi_1 : \rho \to \alpha$ and $\pi_2 : \rho \to \beta$. Then $f_\alpha \circ \pi_1 = f_\beta \circ \pi_2 =$ the unique morphism $\rho \to \gamma$. So $f_\alpha(x) = f_\alpha \circ \pi_1(x, y) = f_\beta \circ \pi_2(x, y) = f_\beta(y)$.
(2) if $f_\alpha(x) = f_\beta(y)$, then similarly to the proof of Theorem 3.3 we get $x \sim f_\alpha(x) = f_\beta(y) \sim y$, so $x \sim y$ follows when bisimilarity is transitive. □

It is natural to ask for intrinsic categorical conditions on the functor T that ensure transitivity of bisimilarity. T is said to *weakly preserve* a pullback square if the T-image of that square satisfies the existence part of the universal property of a pullback, but not necessarily the uniqueness part. Now a pullback of two functions $A \xrightarrow{f} G \xleftarrow{g} B$ is given by the square

$$\begin{tikzcd} R_{fg} \arrow[r, "\pi_B"] \arrow[d, "\pi_A"'] & B \arrow[d, "g"] \\ A \arrow[r, "f"] & G \end{tikzcd}$$

where $R_{fg} = \{(x, y) \in A \times B : f(x) = g(y)\}$ and π_A and π_B are the projections. If f and g are T-morphisms $(A, \alpha) \to (G, \gamma)$ and $(B, \beta) \to (G, \gamma)$, respectively, then Rutten [24, Theorem 4.3] showed that R_{fg} is a bisimulation from α to β *provided that T weakly preserves the pullback of f and g*. The reason why is conveyed by the "cube"

$$
\begin{tikzcd}
R_{fg} \arrow[rr, "\pi_B"] \arrow[dr, "\pi_A"] \arrow[dd, dashed, "\rho"'] & & B \arrow[dr, "g"] \arrow[dd, "\beta"] & \\
& A \arrow[rr, "f"] \arrow[dd, "\alpha"] & & G \arrow[dd, "\gamma"] \\
TR_{fg} \arrow[rr, "T\pi_B"] \arrow[dr, "T\pi_A"'] & & TB \arrow[dr, "Tg"] & \\
& TA \arrow[rr, "Tf"'] & & TG
\end{tikzcd}
$$

Here the top of the cube commutes by definition of R_{fg}, and the base is the T-image of the top. The front and the right side commute because f and g are morphisms. This implies that $Tf \circ (\alpha \circ \pi_A) = Tg \circ (\beta \circ \pi_B)$, so if T weakly preserves the top pullback, then a transition ρ exists as shown to make the left side and the back of the cube commute. This means that the projections are morphisms from (R_{fg}, ρ) to α and β as required.

The pullback of a function with itself is its kernel: $R_{ff} = \mathit{Ker}\, f$. So T is said to *weakly preserve kernels* if it weakly preserves the pullback of any function with itself. This implies that the kernel of any morphism is a bisimulation. In particular it implies that in general ∇_α $(= \mathit{Ker}\, f_{\nabla_\alpha})$ is a bisimulation, hence $\nabla_\alpha = \sim_\alpha$ and so $\sim_\alpha$ is transitive. In other words, weak preservation of kernels ensures transitivity of bisimilarity within each coalgebra. But for transitivity of the global bisimilarity relation it seems that something else is required. We use the condition that T *weakly preserves pullbacks along injective functions*, i.e. it weakly preserves pullbacks of pairs (f, g) at least one of which is injective. It is shown in [6, Theorem 5.7] that this implies that bisimulations between different coalgebras are preserved by subcoalgebras: if (A', α') and (B', β') are subcoalgebras of (A, α) and (B, β), respectively, and R is a bisimulation from α to β, then $R \cap (A' \times B')$ is a bisimulation from α' to β'.

Theorem 3.7. *If T weakly preserves both kernels and pullbacks along injective functions, then T has transitive bisimilarity.*

Proof. Suppose $x \sim_{\alpha\beta} y \sim_{\beta\gamma} z$. We want to show $x \sim_{\alpha\gamma} z$. Let δ be the coproduct (disjoint union) $\alpha + \beta + \gamma$, with insertion morphisms $\iota_\alpha, \iota_\beta, \iota_\gamma$. Then $\iota_\alpha(x) \sim_\delta \iota_\beta(y) \sim_\delta \iota_\gamma(z)$ as bisimilarity is preserved by these morphisms. But as noted above, weak preservation of kernels implies that $\sim_\delta$ is transitive, so $\iota_\alpha(x) \sim_\delta \iota_\gamma(z)$.

Then as T weakly preserves pullbacks along injectives, the restriction R of $\sim_\delta$ to $\mathit{Im}\,\iota_\alpha \times \mathit{Im}\,\iota_\gamma$ is a bisimulation between these subcoalgebras $\mathit{Im}\,\iota_\alpha$ and $\mathit{Im}\,\iota_\gamma$ of δ, with $\iota_\alpha(x)\, R\, \iota_\gamma(z)$. Since these subcoalgebras are isomorphic to α and γ under ι_α and ι_γ, it follows readily that $\{(u, v) : \iota_\alpha(u)\, R\, \iota_\gamma(v)\}$ is a bisimulation from α to γ containing (x, z), so $x \sim_{\alpha\gamma} z$. $\square$

Example 3.8. The covariant powerset functor $\mathcal{P}$ weakly preserves pullbacks, so has transitivity of bisimilarity by 3.7. But there is no final $\mathcal{P}$-coalgebra, for the reasons explained in Example 2.1. $\square$

Examples 3.5 and 3.8 show that transitivity of bisimilarity and the possession of a final coalgebra are independent properties of an endofunctor on **Set**.

4. Abstract logics

We define a *logic* for the functor T to consist of a class Φ and an operation $\models$ that assigns to each T-coalgebra (A, α) a subclass $\models_\alpha$ of $A \times \Phi$. Members of Φ are called *formulas*, and $\models_\alpha$ is the *truth relation* on α. When a pair (x, φ) belongs to $\models_\alpha$ we write $\alpha, x \models \varphi$ and say that the formula φ is *true*, or *satisfied*, at state x in α.

This is a very weak definition of "logic": there is a complete absence of syntactic structure on Φ with corresponding semantic conditions on $\models$. Perhaps "prelogic" would be a better term. But the point is that there are many different kinds of coalgebraic logic, each providing a characterization of bisimilarity in their context, and our aim is to extract what is common to them all. The analysis requires remarkably little common structure, and the weakness of the definition ensures the widespread application of the results that follow from it.

Associated with each pointed coalgebra (α, x) is the "truth-class"

$$\Phi(\alpha, x) = \{\varphi \in \Phi : \alpha, x \models \varphi\}$$

of all formulas true at x in α. A logic is said to have the *Hennessy–Milner (HM) property*, or to be an *HM logic*, if bisimilarity of states is characterized by identity of their truth-classes, which means that for any pointed T-coalgebras (α, x) and (β, y),

$$x \sim_{\alpha\beta} y \quad \text{iff} \quad \Phi(\alpha, x) = \Phi(\beta, y).$$

It is immediate from this definition that the Hennessy–Milner property implies transitivity of T-bisimilarity, and therefore by Theorem 3.3 that coinductive T-coalgebras are simple.

A function $f : A \to B$ between the state sets of α and β is called *truth-invariant* if its action does not alter the truth relation, i.e.

$$\Phi(\alpha, x) = \Phi(\beta, f(x)) \quad \text{for all } x \in A.$$

Note that if f is a morphism, then in general $x \sim_{\alpha\beta} f(x)$, so

(viii) the Hennessy–Milner property implies that every morphism is truth-invariant.

A partial converse to this conclusion is given by the next result, which will be needed to show that a final T-coalgebra exists, and which further demonstrates the strength of the Hennessy–Milner property.

Theorem 4.1. *If a logic has the Hennessy–Milner property, then every truth-invariant function whose codomain is simple must be a morphism.*

Proof. Suppose $f : (A, \alpha) \to (B, \beta)$ is truth-invariant, with β being a simple T-coalgebra. First we find morphisms from α and β into a common *simple* coalgebra γ. To do this, take the coproduct $\alpha + \beta$ with insertion morphisms $\iota_\alpha : \alpha \to \alpha + \beta$ and $\iota_\beta : \beta \to \alpha + \beta$. Then let (G, γ) be the quotient coalgebra $(\alpha + \beta)/\nabla$, where ∇ is the largest congruence of $\alpha + \beta$. Put $f_\alpha = f_\nabla \circ \iota_\alpha$ and $f_\beta = f_\nabla \circ \iota_\beta$, where f_∇ is the quotient morphism, to get T-morphisms $f_\alpha : \alpha \to \gamma$ and $f_\beta : \beta \to \gamma$. γ is simple by Corollary 3.2(1). The situation is depicted in the following diagram.

$$\begin{tikzcd}
A \arrow[rr, "f"] \arrow[dr, "f_\alpha"] \arrow[dd, "\alpha"'] & & B \arrow[dl, "f_\beta"'] \arrow[dd, "\beta"] \\
& G \arrow[dd, "\gamma"] & \\
TA \arrow[rr, "Tf"] \arrow[dr, "Tf_\alpha"'] & & TB \arrow[dl, "Tf_\beta"] \\
& TG &
\end{tikzcd}$$

For f to be a T-morphism requires that $\beta \circ f = Tf \circ \alpha$, i.e. that the “square” in the diagram commutes. We begin by showing that the upper triangle commutes.

Since f_α and f_β are morphisms, they are truth-invariant (viii), as is f by assumption. Thus for any $x \in A$,

$$\Phi(\gamma, f_\alpha(x)) = \Phi(\alpha, x) = \Phi(\beta, f(x)) = \Phi(\gamma, f_\beta(f(x))).$$

Hence $f_a(x) \sim_\gamma f_\beta(f(x))$ by the HM property. But γ is simple, hence coinductive, so then $f_a(x) = f_\beta(f(x))$.

This proves that $f_\alpha = f_\beta \circ f$, so indeed the upper triangle in the diagram commutes. Since T is a functor, the lower triangle then commutes, i.e. $Tf_\alpha = Tf_\beta \circ Tf$. But the fact that f_a and f_β are morphisms means that the left and right “parallelograms” also commute. Given all this commuting, we can chase around the diagram and conclude that

$$Tf_\beta \circ \beta \circ f = Tf_\beta \circ Tf \circ \alpha.$$

If Tf_β is injective, then it can be cancelled from this equation to give the desired result. At this point we invoke the assumption that β is simple to conclude that f_β is injective, by Theorem 3.1(2). Now if $B = \emptyset$, then the presence of f forces $A = \emptyset$, and so as there is only one function $\emptyset \to TB$ the square commutes as desired. Hence we are left with the case $B \neq \emptyset$. But then the injectivity of f_β implies that f_β has a left inverse $g : G \to B$, i.e. $g \circ f_\beta$ is the identity on B. Then Tg is left inverse to Tf_β, from which it follows that Tf_β is injective. (This is just the standard argument that any endofunctor on **Set** preserves injectives with non-empty domain.)

That completes the proof that f is a morphism. $\square$

A logic will be called *small* if its class Φ of formulas is small, i.e. is a set rather than a proper class.

Example 4.2. This is the original example of Hennessy and Milner, providing a small HM logic for the functor $(\mathcal{P}_\omega -)^I$ described in Example 2.1. Here Φ is the set of finitary formulas generated inductively from a propositional constant $\top$ by the standard Boolean connectives together with modalities $\langle i \rangle$ for each $i \in I$. The size of Φ is the maximum of $\aleph_0$ and the size of the set I.

The truth relations are defined by induction of the formation of formulas, with $\alpha, x \models \top$ for all (α, x); the Boolean connectives interpreted as usual; and $\alpha, x \models \langle i \rangle \varphi$ iff $\alpha, y \models \varphi$ for some y with $x \stackrel{i}{\mapsto} y$. The HM property for this logic of $(\mathcal{P}_\omega -)^I$-coalgebras is shown in Theorem 2.2 of [8].

This syntax can be extended by allowing formation of conjunctions of sets of fewer than κ formulas, for some fixed infinite cardinal κ. The result is a small HM logic for the functor $(\mathcal{P}_\kappa -)^I$, where $\mathcal{P}_\kappa A$ is the set of all subsets of A with fewer than κ elements. By allowing conjunctions of *arbitrary* sets of formulas [17], an HM logic for $(\mathcal{P}-)^I$ is obtained. But then Φ becomes a proper class. It turns out that the non-existence of a final $(\mathcal{P}-)^I$-coalgebra implies that there can be no *small* HM logic for $(\mathcal{P}-)^I$. This follows from our next, and main, result. $\square$

Theorem 4.3. *For any functor $T : \mathbf{Set} \to \mathbf{Set}$, the following are equivalent.*

(1) *There exists a small logic for T that has the Hennessy–Milner property.*
(2) *T has a final coalgebra and transitive bisimilarity.*

Proof. First, suppose there is a logic $(\Phi, \models)$ with the properties stated in (1). From the HM property it is immediate that T has transitive bisimilarity.

A final coalgebra will now be constructed as the bisimilarity quotient of a coproduct, an idea that stems from [1,2]. Here we take a logical approach to the coproduct, and also use the fact that in this case the bisimilarity relation on a T-coalgebra is the largest congruence (Theorem 3.3).

Observe that as Φ is a set, the collection of all truth-classes is a set (a subset of the power set of Φ). Hence we can choose a *set* $\mathcal{C}$ of T-coalgebras such that each truth-class (better, truth-*set*) is equal to $\Phi(\alpha, x)$ for some (α, x) with $\alpha \in \mathcal{C}$. Then the coproduct (disjoint union) $\alpha_{\mathcal{C}} = \Sigma\mathcal{C}$ of the members of $\mathcal{C}$ exists as a coalgebra in T-**Coalg**. We can assume that the members of $\mathcal{C}$ are pairwise disjoint, so each $\alpha \in \mathcal{C}$ can be taken to be a subcoalgebra of $\alpha_{\mathcal{C}}$, with the inclusion $\alpha \to \alpha_{\mathcal{C}}$ being a morphism, and therefore truth-invariant by (viii). Hence $\Phi(\alpha, x) = \Phi(\alpha_{\mathcal{C}}, x)$ for any (α, x) with $\alpha \in \mathcal{C}$.

Now let ∇ be the largest congruence on $\alpha_{\mathcal{C}}$, with quotient morphism $f_\nabla : \alpha_{\mathcal{C}} \to \alpha_{\mathcal{C}}/\nabla$. We will show that the quotient $\alpha_{\mathcal{C}}/\nabla$ is a final T-coalgebra. Given any T-coalgebra (B, β), we already know from (vii) in Section 3 that there is at most one T-morphism $g : \beta \to \alpha_C/\nabla$, so all we have to show is that there is at least one such morphism.

For any $y \in B$, choose some (α, x) with $\alpha \in \mathcal{C}$ and $\Phi(\beta, y) = \Phi(\alpha, x)$. Put $g(y) = f_\nabla(x)$. Now $\Phi(\alpha, x) = \Phi(\alpha_{\mathcal{C}}, x)$ (from above), and $\Phi(\alpha_{\mathcal{C}}, x) = \Phi(\alpha_{\mathcal{C}}/\nabla, f_\nabla(x))$ as the morphism f_∇ is truth-invariant. Thus we have $\Phi(\beta, y) = \Phi(\alpha_{\mathcal{C}}/\nabla, g(y))$, i.e. g is truth-invariant. Since $\alpha_{\mathcal{C}}/\nabla$ is simple, Theorem 4.1 then implies that g is a morphism from β to $\alpha_{\mathcal{C}}/\nabla$, completing the proof that $\alpha_{\mathcal{C}}/\nabla$ is final.

[Thus g is uniquely determined, despite the apparent choice in its definition. But if $\Phi(\alpha, x) = \Phi(\alpha', x')$, with $\alpha' \in \mathcal{C}$, then it follows that $\Phi(\alpha_{\mathcal{C}}/\nabla, f_\nabla(x)) = \Phi(\alpha_{\mathcal{C}}/\nabla, f_\nabla(x'))$, and so $f_\nabla(x) = f_\nabla(x')$ by the HM property and the coinductiveness of $\alpha_{\mathcal{C}}/\nabla$.]

For the converse, suppose that T has transitive bisimilarity, and that there is a final coalgebra (G, γ),with f_α being the unique morphism $\alpha \to \gamma$ for each T-coalgebra α. Define a logic by taking $\Phi = G$, so that Φ is small, and putting $\alpha, x \models \varphi$ iff $f_\alpha(x) = \varphi$.

Then in general $\Phi(\alpha, x) = \{f_\alpha(x)\}$, so $\Phi(\alpha, x) = \Phi(\beta, y)$ iff $f_\alpha(x) = f_\beta(y)$. But by Theorem 3.6, $f_\alpha(x) = f_\beta(y)$ iff $x \sim_{\alpha\beta} y$. Thus the HM property holds, and (1) is proved. $\square$

Since the HM property implies transitivity of $\sim$, we immediately get

Corollary 4.4. *If T has a logic with the HM property, then T has a final coalgebra if, and only if, it has a* ***small*** *logic with the HM property.* $\square$

In fact the statement of this corollary is equivalent to that of Theorem 4.3 because, conversely, transitivity of bisimilarity implies the existence of an HM logic for T.[3] To see this, let Φ be the class of all pointed coalgebras, with $\alpha, x \models (\beta, y)$ iff $x \sim_{\alpha\beta} y$. Then $(\beta, y) \in \Phi(\beta, y)$, since $y \sim_{\beta\beta} y$, so if $\Phi(\alpha, x) = \Phi(\beta, y)$ then $(\beta, y) \in \Phi(\alpha, x)$ and so $x \sim_{\alpha\beta} y$. In the reverse direction, if $x \sim_{\alpha\beta} y$ and $\sim$ is transitive, then in general $x \sim_{\alpha\gamma} z$ iff $y \sim_{\beta\gamma} z$, so $\Phi(\alpha, x) = \Phi(\beta, y)$.

The construction of the final coalgebra $\alpha_{\mathcal{C}}/\nabla$ in Theorem 4.3 gives a "syntactic" representation of its states. Each truth-set is equal to some $\Phi(\alpha_{\mathcal{C}}, x)$, and hence to $\Phi(\alpha_{\mathcal{C}}/\nabla,\ f_\nabla(x))$. By the Hennessy–Milner property, distinct states of $\alpha_{\mathcal{C}}/\nabla$ define distinct truth-sets, since the coalgebra is coinductive. Thus the set of states of the final coalgebra corresponds bijectively to the set of all possible truth-sets, and a formula is true at a given state iff it belongs to the corresponding truth-set. That is the basic idea of the canonical model construction: a state determines a truth-set, and in the final coalgebra we can say that a state *is* its truth-set.

5. The intransitive case

If a functor T has transitive bisimilarity, then it has a final coalgebra iff it has small HM logic. But what if bisimilarity is not transitive? Then it is appropriate to consider its transitive closure $\sim^*$. This is the equivalence relation on pointed coalgebras defined by putting $(\alpha, x) \sim^* (\beta, y)$, or $x \sim^*_{\alpha\beta} y$, when there exists an $n \in \omega$ and a sequence $(\alpha, x) = (\alpha_0, x_0), \ldots, (\alpha_n, x_n) = (\beta, y)$ of pointed coalgebras such that $(\alpha_i, x_i) \sim (\alpha_{i+1}, x_{i+1})$ for all $0 \le i < n$.

Lemma 5.1. *If $(\alpha, x) \sim^* (\alpha, y)$, then $x \nabla_\alpha y$.*

Proof. Let $(\alpha, x) = (\alpha_0, x_0), \ldots, (\alpha_n, x_n) = (\alpha, y)$ be a sequence showing $(\alpha, x) \sim^* (\alpha, y)$. Let β be the coproduct $\alpha_0 + \cdots + \alpha_n$ with insertion morphisms $\iota_i : \alpha_i \to \beta$. Let $\iota_\alpha = \iota_0 = \iota_n$. Then

$$\iota_\alpha(x) = \iota_0(x_0) \sim_\beta \iota_1(x_1) \sim_\beta \cdots \sim_\beta \iota_n(x_n) = \iota_\alpha(y)$$

as the insertions preserve bisimilarity. Since $\sim_\beta \subseteq \nabla_\beta$, this gives $\iota_\alpha(x) \nabla_\beta \iota_\alpha(y)$, hence $f_{\nabla_\beta}(\iota_\alpha(x)) = f_{\nabla_\beta}(\iota_\alpha(y))$, so (x, y) belongs to the kernel congruence on α of the morphism $f_{\nabla_\beta} \circ \iota_\alpha$, implying $x \nabla_\alpha y$. $\square$

[3] This observation is due to the referee.

A logic will be said to have the *HM*-property* if, for all pointed T-coalgebras (α, x) and (β, y),

$$x \sim^*_{\alpha\beta} y \quad \text{iff} \quad \Phi(\alpha, x) = \Phi(\beta, y).$$ [4]

If a logic has this weaker property, then morphisms are still truth-invariant, and the conclusion of Theorem 4.1 still holds: if $f : \alpha \to \beta$ is a truth-invariant function with simple codomain β, then f is a morphism. For, in the proof of 4.1, the argument that $\Phi(\gamma, f_\alpha(x)) = \Phi(\gamma, f_\beta(f(x)))$ holds as before, so the HM*-property gives $f_\alpha(x) \sim^*_\gamma f_\beta(f(x))$. But now by Lemma 5.1 and the simplicity of β (i.e. $\nabla_\beta = \Delta_\beta$), this implies $f_\alpha(x) = f_\beta(f(x))$, as required to complete the proof.

Using these facts, a final coalgebra can be constructed from a small logic with the HM*-property by the proof of Theorem 4.3. On the other hand, the construction at the end of that proof produces a small HM*-logic from a final coalgebra. In summary, the statement

T has a final coalgebra iff it has a small HM-logic*

holds for every functor T. Theorem 4.3 is a consequence of this statement.

Note that every functor has a large HM*-logic: let Φ be the class of all pointed coalgebras, with $\alpha, x \models (\beta, y)$ iff $x \sim^*_{\alpha\beta} y$.

6. Conclusion and further questions

An exact relationship has been shown between two quite distinct notions that are fundamental to coalgebraic theory; on the one hand the notion of a final coalgebra, on the other the notion of a logic that characterizes bisimilarity. The proof showed that states of a final coalgebra can be thought of as the truth-sets determined by all states of all coalgebras. The key to the proof was the fact (Theorem 4.1) that a truth-preserving function with a simple codomain must be a coalgebraic morphism.

There remain some questions of interest. Having given (Theorem 3.7) a sufficient functorial condition for transitivity of bisimilarity, an obvious problem is to determine whether it is also *necessary*.

A more substantial question concerns sufficient categorical conditions for the existence of a final coalgebra. It is known [2,3,12,24] that there is a final T-coalgebra whenever T is *bounded*, which means that there is some cardinal number κ such that each state of any T-coalgebra α belongs to some subcoalgebra of α with no more than κ states. Now boundedness does not imply transitivity of bisimilarity, as shown by the functor of Example 3.5, which is bounded with $\kappa = \aleph_0$, and does not have transitivity of bisimilarity, hence has no HM logic. But the logical approach can still be adopted to give another proof that boundedness implies the existence of a final coalgebra: if T is bounded, then there exists a *set* Φ of representatives of the $\sim^*$-equivalence classes of pointed coalgebras, and this gives rise to a small HM*-logic for T. The details are left to the interested reader.

[4] It is readily seen that $x \sim^*_{\alpha\beta} y$ implies $\Phi(\alpha, x) = \Phi(\beta, y)$ for all (α, x) and (β, y) iff $x \sim_{\alpha\beta} y$ implies $\Phi(\alpha, x) = \Phi(\beta, y)$ for all (α, x) and (β, y).

Lastly, noting that a logic has been defined as a language with a *semantics*, we could ask if there is a *proof-theoretic* approach available here. Can we develop an abstract account of proof-relations and deductive consistency that would lead to the construction of final coalgebras whose states were certain consistent sets of formulas closed under suitable proof-relations, as in the classical theory of canonical models?

Acknowledgement

The author thanks the referee for some helpful observations and suggestions.

References

[1] P. Aczel, Non-Well-Founded Sets, in: CSLI Lecture Notes No. 14, CSLI Publications, Stanford University, 1988.
[2] P. Aczel, N. Mendler, A final coalgebra theorem, in: D.H. Pitt et al. (Eds.), Category Theory and Computer Science (Proceedings 1989), in: Lecture Notes in Computer Science, vol. 389, Springer-Verlag, 1989, pp. 357–365.
[3] M. Barr, Terminal coalgebras in well-founded set theory, Theoretical Computer Science 114 (1993) 299–315.
[4] J.A. Bergstra, A. Ponse, S.A. Smolka, Handbook of Process Algebra, Elsevier, 2001.
[5] H.P. Gumm, Elements of the general theory of coalgebras, lUATCS'99, Rand Afrikaans University, Johannesburg, South Africa, 1999, 60 pp. www.Mathematik.uni-marburg.de/~gumm/Papers/publ.html.
[6] H.P. Gumm, T. Schröder, Coalgebraic Structure From Weak Limit Preserving Functors, in: Electronic Notes in Theoretical Computer Science, vol. 33, 2000, www.elsevier.com/locate/entcs.
[7] M. Hennessy, R. Milner, On observing nondeterminism and concurrency, in: J.W. deBakker, J. van Leeuwen (Eds.), Automata, Languages and Programming (Proceedings 1980), in: Lecture Notes in Computer Science, vol. 85, Springer-Verlag, 1980, pp. 299–309.
[8] M. Hennessy, R. Milner, Algebraic laws for nondeterminism and concurrency, Journal of the Association for Computing Machinery 32 (1985) 137–161.
[9] B. Jacobs, Objects and classes, coalgebraically, in: B. Freitag, C.B. Jones, C. Lengauer, H.-J. Schek (Eds.), Object-Orientation with Parallelism and Persistence, Kluwer Academic Publishers, 1996, pp. 83–103.
[10] B. Jacobs, Many-sorted coalgebraic modal logic: a model-theoretic study, Theoretical Informatics and Applications 35 (2001) 31–59.
[11] B. Jacobs, J. Rutten, A tutorial on (co)algebras and (co)induction, Bulletin of the European Association for Theoretical Computer Science 62 (1997) 222–259.
[12] Y. Kawahara, M. Mori, A small final coalgebra theorem, Theoretical Computer Science 233 (2000) 129–145.
[13] A. Kurz, Specifying coalgebras with modal logic, Theoretical Computer Science 260 (2001) 119–138.
[14] J. Lambek, Subequalizers, Canadian Mathematical Bulletin 13 (1970) 337–349.
[15] E.J. Lemmon, An Introduction to Modal Logic, in: K. Segerberg (Ed.), American Philosophical Quarterly Monograph Series, vol. 11, Basil Blackwell, Oxford, 1977 (written in collaboration with Dana Scott).
[16] R. Milner, Calculi for synchrony and asynchrony, Theoretical Computer Science 25 (1983) 267–310.
[17] R. Milner, Communication and Concurrency, Prentice-Hall, 1989.
[18] R. Milner, M. Tofte, Co-induction in relational semantics, Theoretical Computer Science 87 (1991) 209–220.
[19] H. Reichel, An approach to object semantics based on terminal co-algebras, Mathematical Structures in Computer Science 5 (1995) 129–152.
[20] M. Rößiger, From modal logic to terminal coalgebras, Theoretical Computer Science 260 (2001) 209–228.
[21] M. Rößiger, Coalgebras and Modal Logic, in: Electronic Notes in Theoretical Computer Science, vol. 33, 2002, www.elsevier.com/locate/entcs.

[22] J. Rutten, A calculus of transition systems (towards universal coalgebra), in: A. Ponse, M. deRijke, Y. Venema (Eds.), Modal Logic and Process Algebra, in: CSLI Lecture Notes No. 53, CSLI Publications, Stanford, CA, 1995, pp. 231–256.
[23] J. Rutten, Universal coalgebra: a theory of systems, Tech. Rep. CS-R9652, Centrum voor Wiskunde en Informatica (CWI), Amsterdam, 1996.
[24] J. Rutten, Universal coalgebra: a theory of systems, Theoretical Computer Science 249 (1) (2000) 3–80.
[25] J. Rutten, D. Turi, Initial algebra and final coalgebra semantics for concurrency, in: J.W. deBakker, W.-P. deRoever, G. Rozenberg (Eds.), A Decade of Concurrency: Reflections and Perspectives: REX School/Symposium, in: Lecture Notes in Computer Science, vol. 803, Springer Verlag, 1994, pp. 530–882.
[26] K. Segerberg, An Essay in Classical Modal Logic, in: Filosofiska Studier, vol. 13, Uppsala Universitet, 1971.