Coalgebra and Modal Logic: a first introduction

Yde Venema*

©YV 2026

DO NOT DISTRIBUTE

Abstract

These notes give a brief introduction to the theory of universal coalgebra and coalgebraic modal logic.

*Institute for Logic, Language and Computation, University of Amsterdam, Science Park 107, NL–1098XG Amsterdam. E-mail: y.venema@uva.nl.

Contents

1 Introduction **1-1**

2 Final Coalgebras and Coinduction **2-1**

3 Bisimilarity and Behavioural Equivalence **3-1**

4 Covarieties **4-1**

5 Coalgebraic modalities via relation lifting **5-1**

6 Coalgebraic modalities via predicate liftings **6-1**

7 One-step logic **7-1**

8 Soundness and completeness **8-1**

9 Coalgebraic fixpoint logic **9-1**

A Appendix: Basic mathematical definitions **A-1**

B Appendix: The Category Set and its Functors **B-1**

1 Introduction

Starting from concrete examples, this chapter introduces (set-based) coalgebras, together with some of the most important coalgebraic concepts, including coinduction, behavioural equivalence and bisimilarity. We then give a first discussion of the relation between coalgebra and modal logic, and we give some examples of coalgebraic modal logics. In the final section we provide some pointers to the literature on coalgebra and modal logic.

1.1 State-based evolving systems

Example 1.1 Perhaps the simplest example of a computational process is the following black-box machine with two buttons, h and n. If we press the h-button, the machine displays some value from a data set C. No matter how many times we press the h-button, this value remains the same. Each time we push the n-button, however, we may observe a different value by pushing the h-button.

A natural way to formally describe this device is as a set S of internal states (that are not visible to the user), together with two maps

$$\begin{array}{ll} h: & S \to C \\ n: & S \to S, \end{array}$$

where $h(s)$ indicates the observable value at state s, and $n : S \to S$ is a function mapping a state $s \in S$ to its unique next state.

All that we may observe of a state s, is the *stream*

$$h(s) \cdot h(n(s)) \cdot h(n(n(s))) \cdot h(n^3(s)) \cdots$$

of data. This stream will be called the *behaviour* of s in $\mathbb{S} = (S, h, n)$, notation: $\mathsf{beh}_{\mathbb{S}}(s)$.

Two states s and s' in two black boxes can then be called *behaviourally equivalent*, notation: $\mathbb{S}, s \simeq \mathbb{S}', s'$, if they display the same behaviour, that is, if $\mathsf{beh}_{\mathbb{S}}(s) = \mathsf{beh}_{\mathbb{S}'}(s')$.

Example 1.2 A *deterministic finite state automaton* or *DFA* over an alphabet C is a triple $\mathbb{A} = (A, \delta, F)$, where A is a finite set of states, $\delta : A \times C \to A$ is the *transition map* of $\mathbb{A}$, and $F \subseteq A$ is the set of *accepting states* of the automaton. In contrast to the usual presentation we do not take the initial state of a DFA to be part of its structure; rather, we will consider *initialised versions* of DFAs, where we explicitly single out an initial state for the automaton.

Let C^* denote the set of finite words over C, then we may extend δ to a map $\widehat{\delta} : A \times C^* \to A$ as follows:

$$\begin{array}{lcl} \widehat{\delta}(a, \varepsilon) & := & a \\ \widehat{\delta}(a, cw) & := & \widehat{\delta}(\delta(a, c), w). \end{array}$$

We define, for a state $a \in A$,

$$L_{\mathbb{A}}(a) := \{w \in C^* \mid \widehat{\delta}(a, w) \in F\},$$

as the *language* accepted by $\mathbb{A}$, initialized at a. Two initialized automata $(\mathbb{A}, a)$ and $(\mathbb{A}', a')$ are *(language) equivalent* if they accept exactly the same words, that is, if $L_{\mathbb{A}}(a) = L_{\mathbb{A}'}(a')$.

Example 1.3 The key structures featuring in the semantics of *modal logic* are Kripke frames and Kripke models. A *Kripke frame* is a pair (S, R) consisting of a set S of objects called *states*, *points* or *worlds*, and an *accessibility relation* $R \subseteq S \times S$. A *Kripke model* is a triple $\mathbb{S} = (S, R, V)$ such that (S, R) is a Kripke frame (the *underlying Kripke frame* of the model), and V is a *valuation*, i.e., a map $\mathsf{Q} \to PS$, where Q is some fixed set of *proposition letters*.

A *bisimulation* between two Kripke models $\mathbb{S}$ and $\mathbb{S}'$ is a binary relation $Z \subseteq S \times S'$ such that, for all $s \in S$ and $s' \in S'$ with Zss' the following conditions hold:

(atom) s and s' satisfy the same proposition letters;

(forth) for all $t \in S$ such that Rst there is a $t' \in S'$ with $R's't'$ and Ztt';

(back) for all $t' \in S'$ such that $R's't'$ there is a $t \in S$ with Rst and Ztt'.

If there is a bisimulation Z linking s and s' we say that s and s' are *bisimilar*, notation $\mathbb{S}, s \leftrightarrows \mathbb{S}', s'$ (or $Z : \mathbb{S}, s \leftrightarrows \mathbb{S}', s'$ if we want to make the bisimulation explicit).

Given a modal language L, we let $Th^L_{\mathbb{S}}(s)$ denote the set of L-formulas that are true at s in $\mathbb{S}$; then two states are called *(L-)equivalent* (notation: $\equiv_L$) if they satisfy the same L-formulas. A theme in the theory of modal logic is to study the relation between equivalence and bisimilarity. A class of models C is said to have the *Hennessy-Milner* property with respect to a language L if $\equiv_L \;=\; \leftrightarrows$ on C.

Example 1.4 The theory of *non-well-founded sets* provides an alternative to the standard axiomatic set theories by allowing sets to contain themselves, or otherwise violate the rule of well-foundedness. More in detail, in non-well-founded set theories, the Foundation Axiom *FA* is replaced by axioms implying its negation. For instance, working with the *anti-foundation axiom AFA* we may associate, with each so-called *apg* or *accessible pointed graph* (that is, a directed graph such that every node can be reached via a finite path from a specified *root* of the graph) a *hyperset*, that is, a set that is not necessarily well-founded. And, two apgs yield the *same* set iff they are bisimilar.

Example 1.5 As a final example of a state-based evolving system we mention Markov chains: transition systems that evolve probabilistically. Recall that a *(discrete) probability distribution* on a set S is a map $\mu : S \to [0, 1]$ such that $\sum_{s \in S} \mu(s) = 1$. Formally, a Markov chain can be modelled as a pair (S, σ), where σ assigns a probability distribution σ_s to each state s.

For a concrete example, think of a gambler wagering €1 on a series of fair coin tosses – this series may be indefinite, or end if the gambler loses his money. This experiment can be modelled by the Markov chain (S, σ) where $S = \{s_n \mid n \in \omega\}$, with state s_n representing the state where the gambler owns €n. For $n > 0$ we have that σ_{s_n} assigns a 0.5 probability to both s_{n-1} and s_{n+1} (and a probability 0 to all other states), while σ_{s_0} assigns a 1.0 probability to s_0 (and a probability 0 to all other states).

1.2 Coalgebras and their morphisms

As we will see now, the structures described in the previous section all are specimens of *coalgebras*. Universal Coalgebra is a theory of state-based evolving systems, formulated in the language of category theory.[1]

[1] See the appendix for some background definitions on category theory.

Definition 1.6 Given an endofunctor $T : \mathsf{C} \to \mathsf{C}$ on some category C, a T-coalgebra is a pair (X, ξ) where X is an object in C and $\xi : X \to TX$ is an arrow in C. We will sometimes refer to T as the *type* of (X, ξ). If, for an arrow $f : X' \to X$, the following diagram commutes:

$$\begin{array}{ccc} X' & \xrightarrow{f} & X \\ {\scriptstyle \xi'}\downarrow & & \downarrow{\scriptstyle \xi} \\ TX' & \xrightarrow{Tf} & TX \end{array} \tag{1}$$

we call f a *(coalgebra) morphism* from $\mathbb{X}' = (X', \xi')$ to $\mathbb{X} = (X, \xi)$, and write $f : \mathbb{X}' \to \mathbb{X}$.

We let $\mathsf{Coalg}_{\mathsf{C}}(T)$ denote the category with T-coalgebras as objects and T-coalgebra morphisms as arrows; the category C will be called the *base* category of $\mathsf{Coalg}_{\mathsf{C}}(T)$. $\lhd$

We will usually (but not always) confine our attention to *systems*, that is, coalgebras over the category Set. Intuitively, a set functor T specifies the one-step dynamics that a system can engage in.

Definition 1.7 A *set functor* is an endofunctor $T : \mathsf{Set} \to \mathsf{Set}$ on the category Set of sets and functions; given such a set functor T, we will sometimes refer to T-coalgebras as T*-systems*. Given the prominence of these coalgebras we will often simply write $\mathsf{Coalg}(T)$ rather than $\mathsf{Coalg}_{\mathsf{Set}}(T)$ to denote the class of all T-systems. Where $\mathbb{S} = (S, \sigma)$ is a T-system, we refer to S and σ as, respectively, the *carrier* or *state space* and the *transition map* or *coalgebra map* of $\mathbb{S}$. A *pointed* or *initialized* T-system is a triple (S, σ, s) such that (S, σ) is a T-system and $s \in S$. $\lhd$

The coalgebraic viewpoint on systems combines wide applicability and mathematical simplicity: since every set functor determines its own type of coalgebra, notions, properties and results of state-based systems can be uniformly explained just in terms of properties of their type functors. This applies to systems as diverse as streams, probabilistic transition systems, automata, Kripke structures and neighbourhood frames. In the appendix we give a list of set functors; here we give a few examples of the associated coalgebras.

Example 1.8 (a) The black boxes of Example 1.1 are systems of the functor type $K_C \times Id$, where K_C is the constant functor associated with the set C, and Id is the identity functor on Set.

(b) Deterministic finite automata (Example 1.2) are systems of type $2 \times Id^C$ — note that here we write 2 rather than K_2 to denote the constant functor associated with the set $2 = \{0, 1\}$. To see this, consider a coalgebra $\mathbb{X} = (X, \xi)$ of this type; then ξ determines, for each state $x \in X$, two things: an element $\xi_0(x)$ of the set 2, specifying whether x is accepting ($\xi_0(x) = 1$) or not ($\xi_0(x) = 0$), and an element $\xi_1(x) \in X^C$, that is, a map $\xi_1(x) : C \to X$ providing a successor of x in X for each letter $c \in C$. A $2 \times Id^C$-coalgebra (X, ξ) thus corresponds to the DFA $X, \xi_0^{-1}(1), \xi_1)$.

(c) Kripke frames are coalgebras for the powerset functor P, whereas Kripke models are coalgebras for the functor $K_{P\mathsf{Q}} \times P$. We leave it for the reader to verify that, given two Kripke

frames $\mathbb{S}$ and $\mathbb{S}'$, a map $f : S \to S'$ is a coalgebra morphism iff it is a bounded morphism, that is, for every $s \in S$ it satisfies:

$\circ$ if $(s,t) \in R$ then $(fs, ft) \in R'$;

$\circ$ if $(fs, t') \in R'$ then $(s,t) \in R$ for some t such that $ft = t'$.

(d) Non-well-founded sets can be represented by certain pointed P-systems.

(e) Markov chains are D-systems, where D is the distribution functor (which assigns, to a set S, a discrete probability distribution on S).

(f) For every set functor T we will allow the empty T-coalgebra $(\varnothing, \varnothing)$.

1.3 Final coalgebras and coinduction

For many coalgebra types T one may associate with an arbitrary state s in an arbitrary T-coalgebra $\mathbb{S}$, a natural notion of *behaviour*. This can often be formalised by defining a *behaviour map* and proving that this map is the unique coalgebra morphism from $\mathbb{S}$ to some *final* or *terminal* coalgebra $\mathbb{Z}$ of type T.

Definition 1.9 Let T be an endofunctor on some category C. A *T-coalgebra* $\mathbb{Z} = (Z, \zeta)$ is *final* or *terminal* if it is a final object in the category $\mathsf{Coalg}_{\mathsf{C}}(T)$; that is, if for every T-coalgebra $\mathbb{X} = (X, \xi)$ there is a *unique* morphism from $\mathbb{X}$ to $\mathbb{Z}$; this morphism will be denoted as $\mathsf{beh}_{\mathbb{X}} : \mathbb{X} \to \mathbb{Z}$. $\lhd$

Note that final coalgebras, when they exist, are unique modulo isomorphism. For this reason we will often speak of *the* final T-coalgebra of a functor T.

Example 1.10 A *stream* over a set C is a map $\alpha : \omega \to C$ (where ω is the set of natural numbers). We may turn the set C^ω of C-streams into a $K_C \times Id$-coalgebra itself by endowing it with a coalgebra map $\gamma := (\mathsf{h}, \mathsf{t}) : C^\omega \to C \times C^\omega$. Here we define the maps $\mathsf{h} : C^\omega \to C$ and $\mathsf{t} : C^\omega \to C^\omega$ by putting, for an arbitrary C-stream $\alpha : \omega \to C$:

$$\begin{array}{lcl} \mathsf{h}(\alpha) & := & \alpha(0) \\ \mathsf{t}(\alpha) & := & \lambda n.\alpha(n+1). \end{array}$$

That is, the coalgebra map γ splits an infinite C-stream $c_0c_1c_2\cdots$ into its *head* c_0 and its *tail* $c_1c_2c_3\cdots$.

It is then not very hard to prove that the *stream coalgebra* (C^ω, γ) is a final coalgebra for the functor $K_C \times Id$: this boils down to showing that, for an arbitrary 'black box machine' $\mathbb{S} = (S, h, n)$, the behaviour map $\mathsf{beh} : S \to C^\omega$ is the *unique* coalgebra morphism from $\mathbb{S}$ to (C^ω, γ).

Example 1.11 For a second example, fix an alphabet C and define a (C-)*language* to be any set of finite words. Further on we will see that we can endow the collection $P(C^*)$ of C-languages with a very natural coalgebra map for the functor $2 \times Id^C$ of deterministic finite C-automata, and prove that the resulting structure is in fact a final $2 \times Id^C$-coalgebra.

Finality is also the key categorical concept underlying the important coalgebraic principle of *coinduction*. Here is a first example.

Example 1.12 Take the function zip that merges two streams by taking elements from either stream in turn. For a coalgebraic definition of this map, define the transition map $\delta : (C^\omega \times C^\omega) \to C \times (C^\omega \times C^\omega)$ as follows:

$$\delta(\alpha, \beta) := (\mathsf{h}(\alpha), (\beta, \mathsf{t}(\alpha)),$$

where h and t are the maps defined in Example 1.10. This defines a $K_C \times Id$-coalgebra on the set $(C^\omega \times C^\omega)$, so that by finality of the stream coalgebra (C^ω, γ) there is a (unique) map $\mathsf{zip} : C^\omega \times C^\omega \to C^\omega$ which is a coalgebra morphism from $(C^\omega \times C^\omega, \delta)$ to $(C^\omega, \langle \mathsf{h}, \mathsf{t} \rangle)$:

$$\begin{array}{ccc} C^\omega \times C^\omega & \xrightarrow{\quad\mathsf{zip}\quad} & C^\omega \\ \downarrow{\scriptstyle \delta} & & \downarrow{\scriptstyle \langle \mathsf{h},\mathsf{t} \rangle} \\ C \times (C^\omega \times C^\omega) & \xrightarrow[(\mathsf{id}_C, \mathsf{zip})]{} & C \times C^\omega \end{array}$$

One may verify that this coalgebra morphism indeed defines the map that zips two streams together.

Unfortunately, as we will see further on, final coalgebras do not exist for every functor.

Example 1.13 In the categories of Kripke frames and Kripke models, final objects do not exist. The *canonical model* comes close, but to turn this structure into a final coalgebra, we have to enrich the base category Set with topological structure. As a result that we will discuss later on, we may see the canonical *general* frame as a final coalgebra for a suitable base category and coalgebra functor.

1.4 Behavioural equivalence and bisimilarity

Probably the most intuitive notion of equivalence between systems is that of *behavioral*, or *observational*, equivalence. The idea here is to consider two states to be similar if we cannot distinguish them by observations, because they display the same behavior. For instance, we call two deterministic automata (pointed $2 \times Id^C$-coalgebras) equivalent if they recognize the same language. In case the functor T admits a final coalgebra $\mathbb{Z}$, this idea is easily formalized by making state s in coalgebra $\mathbb{S}$ equivalent to state s' in coalgebra $\mathbb{S}'$ if $\mathsf{beh}_{\mathbb{S}}(s) = \mathsf{beh}_{\mathbb{S}'}(s')$. In case the functor does not admit a final coalgebra, we generalize this demand as follows.

Definition 1.14 Let $\mathbb{S} = (S, \sigma)$ and $\mathbb{S}' = (S', \sigma')$ be two systems for the set functor T. Then $s \in S$ and $s' \in S'$ are *behaviorally equivalent*, notation: $\mathbb{S}, s \simeq_T \mathbb{S}', s'$ if there is a T-system $\mathbb{X} = (X, \xi)$ and homomorphisms $f : \mathbb{S} \to \mathbb{X}$ and $f' : \mathbb{S}' \to \mathbb{X}$ such that $f(s) = f'(s')$. $\lhd$

Remark 1.15 It is easily checked that in case T admits a final coalgebra $\mathbb{Z}$, then indeed $\mathbb{S}, s \simeq_T \mathbb{S}', s'$ iff $\mathsf{beh}_{\mathbb{S}}(s) = \mathsf{beh}_{\mathbb{S}'}(s')$. The direction from right to left is trivial, so assume that s and s' are behaviorally equivalent because of the existence of $\mathbb{X}$, f and f' as in the formulation of the definition. Observe that the map $\mathsf{beh}_{\mathbb{X}} \circ f$ is a coalgebra morphism from $\mathbb{S}$ to $\mathbb{Z}$, and likewise for $\mathsf{beh}_{\mathbb{X}} \circ f'$ and $\mathbb{S}'$. It then follows from the finality of $\mathbb{Z}$ that $\mathsf{beh}_{\mathbb{S}} = \mathsf{beh}_{\mathbb{X}} \circ f$ and $\mathsf{beh}_{\mathbb{S}'} = \mathsf{beh}_{\mathbb{X}} \circ f'$. Hence, from $f(s) = f'(s')$ it follows that $\mathsf{beh}_{\mathbb{S}}(s) = \mathsf{beh}_{\mathbb{S}'}(s')$, as required.

As we will see further on, in many cases of interest, behavioral equivalence can be characterized via the equally fundamental concept of *bisimilarity*, which involves the notion of a *coalgebraic bisimulation*.

Definition 1.16 A *bisimulation* between two systems $(\mathbb{S}, \sigma)$ and $(\mathbb{S}', \sigma')$ is a relation $B \subseteq S \times S'$ for which, as in the diagram below,

$$\begin{array}{ccccc} S & \xleftarrow{\pi} & B & \xrightarrow{\pi'} & S' \\ \downarrow{\scriptstyle \sigma} & & \vdots{\scriptstyle \beta} & & \downarrow{\scriptstyle \sigma'} \\ TS & \xleftarrow[T\pi]{} & TB & \xrightarrow[T\pi']{} & TS' \end{array} \tag{2}$$

there is a (not necessarily unique) coalgebra map $\beta : B \to TB$ such that the two projection maps, π and π', from B to S and S' respectively, are both coalgebra morphisms.

We write $\mathbb{S}, s \leftrightarrows_T \mathbb{S}', s'$ to denote that there is some bisimulation B between $\mathbb{S}$ and $\mathbb{S}'$ that links the states s in $\mathbb{S}$ and s' in $\mathbb{S}'$. We often omit the functor T if this is not likely to cause confusion. $\lhd$

That is, a bisimulation is a relation that, seen as a set, can itself be endowed with a coalgebra structure satisfying some natural conditions.

1.5 Coalgebra and modal logic

Logic comes in when we want to design specification languages for describing the behaviour of state-based evolving systems, and derivation systems for reasoning about this behaviour.

Definition 1.17 An *(abstract) logic* for T-coalgebras is a pair $(L, \Vdash)$ consisting of a set L of *formulas* and, for each T-coalgebra $\mathbb{S} = (S, \sigma)$, a *satisfaction* relation $\Vdash_{\mathbb{S}} \subseteq S \times L$. In case $(s, \varphi) \in \Vdash_{\mathbb{S}}$, we say that φ *is true* or *holds* at $s \in \mathbb{S}$, or that s *satisfies* φ in $\mathbb{S}$; we often write $\mathbb{S}, s \Vdash s$ or even $s \Vdash \varphi$ instead of $s \Vdash_{\mathbb{S}} \varphi$.

Given a state s in a coalgebra $\mathbb{S}$, we define $Th^L_{\mathbb{S}}(s) := \{\varphi \in L \mid \mathbb{S}, s \Vdash \varphi\}$. Conversely, given a formula $\varphi \in L$ and a coalgebra $\mathbb{S}$, we let $[\![\varphi]\!]^{\mathbb{S}}$ denote the set of states in $\mathbb{S}$ where φ holds, that is, $[\![\varphi]\!]^{\mathbb{S}} := \{s \in S \mid \mathbb{S}, s \Vdash \varphi\}$. If $Th^L_{\mathbb{S}}(s) = Th^L_{\mathbb{S}'}(s')$, we say that s and s' are *L-equivalent*, and we write $\mathbb{S}, s \equiv_L \mathbb{S}', s'$ (or simply $s \equiv_L s'$ if $\mathbb{S}$ and $\mathbb{S}'$ are understood).

Finally, we will call a formula *satisfiable* if it is satisfied at some state in some coalgebra, and *valid* if it holds at every state in every coalgebra. $\lhd$

In the same way that universal coalgebra tries to give an account of state-based evolving systems, *uniformly* in the coalgebra type T, research in coalgebraic logic has been directed towards a development of logical languages and derivation systems that are similarly uniform in the parameter T. Apart from uniformity, here are some other desiderata for a coalgebraic logic.

Definition 1.18 Let $(L, \Vdash)$ be a logic for a coalgebra type T. We say that this logic is *invariant (for behavioural equivalence)* if $\mathbb{S}, s \simeq \mathbb{S}', s'$ implies $\mathbb{S}, s \equiv_L \mathbb{S}', s'$ and *expressive* if conversely, $\mathbb{S}, s \equiv_L \mathbb{S}', s'$ implies that s and s' are behaviourally equivalent. The logic is *decidable* if there is an algorithm that decides, on input $\varphi \in L$, whether there is some pointed coalgebra satisfying φ. $\lhd$

In addition, for practical purposes one generally wants the logic to be *finitary* in the sense that formulas are finite objects. Other desirable properties of a coalgebraic logic include good model-theoretic behaviour, and the existence of a derivation system that is sound and complete for the collection of valid formulas.

With Kripke models as paradigmatic examples of coalgebra, and modal logic providing *the* bisimulation-invariant logic for Kripke models, it should come as no surprise that most coalgebraic logics can be seen as generalisations of basic modal logic in some sense. The literature on coalgebra witnesses different ways to generalise basic modal logic from Kripke structures to arbitrary systems; here we mention two approaches.

First, however, we briefly discuss the role of *proposition letters* in coalgebraic modal logic. Generalising the relation between Kripke models and Kripke frames, we introduce the notion of a T-model.

Definition 1.19 Let T be a set functor, and let Q be an arbitrary but fixed set of proposition letters. A *T-model* is a triple (S, σ, V) such that (S, σ) is a T-coalgebra, and $V : \mathsf{Q} \to PS$ is a *valuation*. A morphism between T-models $\mathbb{S} = (S, \sigma, V)$ and $\mathbb{S}' = (S', \sigma', V')$ is a coalgebra morphism $f : (S, \sigma) \to (S, \sigma')$ such that $s \in V(p)$ iff $fs \in V'(p)$, for all $s \in S$. $\lhd$

Remark 1.20 There are two natural ways to think about T-models: either as T-coalgebras extended with a Q-valuation, or as coalgebras for the functor $T_{\mathsf{Q}} := K_{P\mathsf{Q}} \times T$. (Clearly, in the latter case it would be more natural to represent the valuation V as its associated *colouring* $V^\flat : S \to P\mathsf{Q}$ given by $V^\flat(s) := \{p \in \mathsf{Q} \mid s \in V(p)\}$, cf. Example A.5.) In these notes we will generally take the first perspective, since it is more compatible with the perspective on proposition letters as *variables*. Nevertheless we will apply various coalgebraic definitions to T-models as if they were indeed $K_{P\mathsf{Q}} \times T$-coalgebras.

The notion of a morphism between T-models is already an example of this: it is easy to see that morphism between two T-models is nothing but a coalgebra morphism between the associated T_{Q}-coalgebras.

Similarly, we will say that a relation $Z \subseteq S \times S'$ is a *bisimulation* between the two *T-models* $\mathbb{S} = (S, \sigma, V)$ and $\mathbb{S}', \sigma', V'$ if it is a bisimulation between $\mathbb{S}$ and $\mathbb{S}'$, seen as two T_{Q}-coalgebras. It is easy to verify that this holds iff Z is a T-bisimulation between (S, σ) and (S', σ') that in addition satisfies the condition (atom) that Z-related states satisfy the same proposition letters.

Let us now take a quick look at two distinct approaches towards coalgebraic modal logic.

Example 1.21 In the first approach towards coalgebraic modal logic, which is completely parametric in the functor T, the set of formulas L is closed under the following clause, which introduces a modal operator ∇:

$$\text{if } \alpha \in TX \text{ for some finite set } X \text{ of formulas, then } \nabla\alpha \text{ is a formula.}$$

In the case of the powerset functor ($T = P$), we can write, for instance, $\nabla\{\varphi_0, \varphi_1\}$, where φ_0 and φ_1 are formulas. The formula $\nabla\{\varphi_0, \varphi_1\}$ will be equivalent to $(\Diamond\varphi_0 \wedge \Diamond\varphi_1) \wedge \Box(\varphi_0 \vee \varphi_1)$. In general, the semantics of ∇ in a Kripke structure $\mathbb{S} = (S, \sigma)$ will be given as

$$\mathbb{S}, s \Vdash \nabla\alpha \text{ iff } (\sigma(s), \alpha) \in \overline{P}(\Vdash),$$

where $\overline{P}(\Vdash) \subseteq PS \times PL$ is the (Egli-Milner) *lifting*[2] of the binary satisfaction relation $\Vdash$.

As we will see, this approach generalises well to any set functor T that 'preserves weak pullbacks' — the point of this condition being that T preserves weak pullbacks iff its lifting $\overline{T}$ preserves relation composition.

The ∇-logic provided by the relation-lifting approach described in Example 1.21 may provide coalgebraic logics in a completely uniform way, but its unusual syntax makes it not easy to work with.

The second approach to coalgebraic logic provides coalgebraic logics with a more standard modal syntax. Here, the modalities of the language correspond to so-called *predicate liftings*, where an n-ary predicate lifting is a natural transformation $\breve{P}^n \dot{\to} \breve{P}T$. Here we confine ourselves to a few examples of such coalgebraic modalities.

Example 1.22 The standard interpretation of the modalities $\Diamond$ and $\Box$ in Kripke structures can be formulated as follows:

$$\begin{array}{lll} \mathbb{S}, s \Vdash \Box\varphi & \text{iff} & R(s) \subseteq [\![\varphi]\!] \\ \mathbb{S}, s \Vdash \Diamond\varphi & \text{iff} & R(s) \cap [\![\varphi]\!] \neq \varnothing. \end{array}$$

Example 1.23 *Monotone modal logic* is a variant of standard modal logic where formulas are interpreted in so-called monotone neighbourhood models. These are structures of the form $\mathbb{S} = (S, \sigma, V)$ where S is a set of states, V is a valuation, and σ is a map $S \to PPS$ that assigns to each state $s \in S$ a collection $\sigma(s) \subseteq PS$ of neighbourhoods. Here, each collection $\sigma(s)$ is required to be *upwards closed* in the sense that $X \in \sigma(s)$ implies $Y \in \sigma(s)$ for all Y with $X \subseteq Y \subseteq S$.

In these structures we may interpret the modalities $\Diamond$ and $\Box$ as follows:

$$\begin{array}{lll} \mathbb{S}, s \Vdash \Box\varphi & \text{iff} & [\![\varphi]\!] \in \sigma(s) \\ \mathbb{S}, s \Vdash \Diamond\varphi & \text{iff} & (S \setminus [\![\varphi]\!]) \notin \sigma(s). \end{array}$$

Using the upwards-closedness of $\sigma(s) \subseteq P(S)$ it is not hard to show that $\Box\varphi$ holds at s iff s has a neighbourhood $U \in \sigma(s)$ such that $\mathbb{S}, u \Vdash \varphi$, for each $u \in U$, whereas $\Diamond\varphi$ holds at s iff every neighbourhood $U \in \sigma(s)$ contains some point u where φ holds.

To see how monotone modal logic generalises standard modal logic, think of a Kripke model (S, R, V) as the neighbourhood model $(S, \widehat{\sigma}, V)$, where $\widehat{\sigma}(s) := \{X \in PS \mid R(s) \subseteq X\}$.

[2]The Egli-Milner lifting of a relation $R \subseteq S \times S'$ is the relation $\overline{P}(R) \subseteq PS \times PS'$ given by $(X, X') \in \overline{P}(R)$ iff for all $x \in X$ there is an $x' \in X'$ such that Rxx' and for all $x' \in X'$ there is an $x \in X$ such that Rxx'.

Example 1.24 Let (S, σ) be a Markov chain, that is, a coalgebra for the distribution functor D. Given a rational number $q \in [0, 1]$, we introduce a modality $\Diamond_q$, with the following intended meaning:

$$\mathbb{S}, s \Vdash \Diamond_q \varphi \text{ iff } \sum_{t \in [\![\varphi]\!]^{\mathbb{S}}} \mu_s(t) > q.$$

That is, the formula $\Diamond_q \varphi$ holds at s iff the probability that φ holds at the next state after s is bigger than q.

1.6 Literature

Here are some relevant texts on coalgebra and modal logic. First we mention some books:

- J. Barwise and L. Moss, *Vicious Circles*, CSLI Publications, 1996.
- B. Jacobs, *Introduction to Coalgebra: towards mathematics of states and observation*, Cambridge University Press, 2016.
- J. Rutten, *The Method of Coalgebra: exercises in coinduction*, CWI, Amsterdam, The Netherlands, 2019, ISBN 978-90-6196-568-8.

Here is a list of introductory and survey articles:

- B. Jacobs and J. Rutten, *A tutorial on (co)algebras and (co)induction*, Bulletin of the European Association for Theoretical Computer Science, 62 (1997), pp. 222–259..
- J.J.M.M Rutten, *Universal coalgebra: a theory of systems*, Theoretical Computer Science 249 (2000), pp. 3-80.
- H.P. Gumm, *Universelle coalgebra*, appendix to Th. Ihringer: Universelle Algebra, Heldermann Verlag, Berlin, 2003. (available from the web page of H.P. Gumm.)
- A. Kurz, A. Palmigiano and Y. Venema, *Coalgebra and logic: an overview*, Journal of Logic and Computation 20 (2010), pp. 985-988.
- C. Kupke and D. Pattinson, *Coalgebraic Semantics of Modal Logics: an Overview*, Theoretical Computer Science, 412(38) (2011), pp. 5070-5094.
- C. Cîrstea, A. Kurz, D. Pattinson, L. Schröder and Y. Venema, *Modal logics are coalgebraic*, The Computer Journal, 54 (2011), pp. 31-41.

2 Final Coalgebras and Coinduction

In section 1.3 we introduced final coalgebras. In this chapter we study the concept in more detail, and we see how it relates to the fundamental coalgebraic definition and proof principle of *coinduction*.

As a first example of a final coalgebra, it is instructive to look at a base category different from Set.

Example 2.1 Let $\mathbb{C} = (C, \leq)$ be an arbitrary poset, that is, $\leq$ is a reflexive, transitive and antisymmetric relation on the set C. We may think of $\mathbb{C}$ as a category by taking C as the set of objects and providing a unique arrow between any pair of elements $c, d \in C$ for which $c \leq d$. An *endofunctor* on $\mathbb{C}$ is then nothing but a *monotone* or *order-preserving* function $F : C \to C$. Given that arrows between objects are unique if they exist, a *coalgebra* $(c, \gamma : c \to Fc)$ for such a functor can be identified with its carrier c, and conversely, any $c \in C$ for which $c \leq Fc$ is the carrier of an F-coalgebra. In other words, we may identify F-coalgebras with the *prefixpoints* of F.

We leave it for the reader to verify that a final coalgebra for a functor F on $\mathbb{C}$ is a *greatest fixpoint* of the map F.

2.1 The language coalgebra

As a key example of a final coalgebra we will show how to endow the collection of *languages* over some finite alphabet C with coalgebra structure that turns it into the final coalgebra for the set functor $2 \times Id^C$ associated with deterministic automata.

Here we will represent a deterministic automaton over an alphabet C as a triple $\mathbb{S} = (S, \chi, \tau)$, where $\chi : S \to 2$ and $\tau : S \to S^C$ correspond to the acceptance condition and the transition map, respectively. Note that we drop the condition that the carrier of the automaton is finite. We will also use the convention that $s \xrightarrow{a} t$ means $t = \tau(s)(a)$ and $s{\downarrow}$ indicates that s is accepting, i.e., $\chi(s) = 1$.

As we saw in the introduction, we can identify deterministic automata with coalgebras of the functor $2 \times Id^C$. It is easy to see that a map $f : S \to S'$ is a coalgebra morphism between two automata $\mathbb{S} = (S, \chi, \tau)$ and $\mathbb{S}' = (S', \chi', \tau')$ if it satisfies, for all $s \in S$ and $a \in C$, the conditions $\chi(s) = \chi'(fs)$ and $f(\tau(s)(a)) = \tau'(fs)(a)$.

Definition 2.2 Consider the following *language coalgebra* $\mathbb{L} := (\mathcal{L}_C, \omega, \delta)$, where
- $\mathcal{L}_C := P(C^*)$ is the collection of all *languages* over C,
- $\omega : \mathcal{L}_C \to 2$ is given by $\omega(L) := 1$ if $\varepsilon \in L$, and $\omega(L) = 0$ if $\varepsilon \notin L$;
- $\delta : \mathcal{L}_C \to (\mathcal{L}_C)^C$ is the map given by $\delta(L)(a) := L_a$, the so-called *a-derivative* of L:

$$L_a := \{u \in C^* \mid au \in L\}.$$

If no confusion concerning the alphabet is likely, we will usually write $\mathcal{L}$ rather than $\mathcal{L}_C$. $\lhd$

Recall that for an arbitrary automaton $\mathbb{S} = (S, \chi, \tau)$, we defined the language recognized by a state $s \in S$ by putting

$$L_{\mathbb{S}}(s) := \{u \in C^* \mid \chi(\widehat{\tau}(s)(u)) = 1\},$$

where $\widehat{\tau} : S \to S^{C^*}$ is inductively defined by putting $\widehat{\tau}(s)(\varepsilon) := s$ and $\widehat{\tau}(s)(cu) := \widehat{\tau}\big(\tau(s)(c)\big)(u)$.

We claim that, for any alphabet C, $\mathbb{L}$ is the final coalgebra of type $2 \times Id^C$, with the language maps as the witnessing coalgebra morphisms.

Proposition 2.3 (Finality of $\mathbb{L}$) *For any $2 \times Id^C$-coalgebra $\mathbb{S}$, the map $L_{\mathbb{S}}$ is the unique coalgebra morphism $L_{\mathbb{S}} : \mathbb{S} \to \mathbb{L}$.*

Proof. Fix $\mathbb{S} = (S, \chi, \tau)$. We first show that $L_{\mathbb{S}}$ is a coalgebra morphism. For acceptance, we check that $\chi(s) = \omega(L_{\mathbb{S}}(s))$:

$$\begin{aligned} \chi(s) = 1 \text{ iff } & \varepsilon \in L_{\mathbb{S}}(s) && \text{(definition } L_{\mathbb{S}}\text{)} \\ \text{iff } & \omega(L_{\mathbb{S}}(s)) = 1 && \text{(definition } \omega\text{)} \end{aligned}$$

With respect to the transition function, we need to show that $L_{\mathbb{S}}(\tau(s)(c)) = \delta(L_{\mathbb{S}}(s))(c)$, for all $s \in S$ and $c \in C$. But this identity holds because of the following chain of equivalences, for an arbitrary word $u \in C^*$:

$$\begin{aligned} u \in L_{\mathbb{S}}(\tau(s)(c)) \text{ iff } & \chi\big(\widehat{\tau}(\tau(s)(c))(u)\big) = 1 && \text{(definition } L_{\mathbb{S}}\text{)} \\ \text{iff } & \chi(\widehat{\tau}(s)(cu))) = 1 && \text{(definition } \widehat{\tau}\text{)} \\ \text{iff } & cu \in L_{\mathbb{S}}(s) && \text{(definition } L_{\mathbb{S}}\text{)} \\ \text{iff } & u \in \delta(L_{\mathbb{S}}(s))(c) && \text{(definition } \delta\text{)} \end{aligned}$$

Second, we prove uniqueness. Assuming that $f : \mathbb{S} \to \mathbb{L}$ is a coalgebra morphism, we need to show that $f = L_{\mathbb{S}}$. It suffices to show that any word $u \in C^*$ satisfies the following:

$$\text{for all } s \in S : \quad u \in L_{\mathbb{S}}(s) \text{ iff } u \in f(s). \tag{3}$$

We will prove (3) by induction on u. In the base case, where $u = \varepsilon$, we have

$$\begin{aligned} \varepsilon \in L_{\mathbb{S}}(s) \text{ iff } & \omega(L_{\mathbb{S}}(s)) = 1 && \text{(definition } \omega\text{)} \\ \text{iff } & \chi(s) = 1 && (L_{\mathbb{S}} \text{ is a morphism}) \\ \text{iff } & \omega(f(s)) = 1 && (f \text{ is a morphism}) \\ \text{iff } & \varepsilon \in f(s) && \text{(definition } \omega\text{)} \end{aligned}$$

Now assume that $u = cv$, for some $c \in C$, then we find

$$\begin{aligned} cv \in L_{\mathbb{S}}(s) \text{ iff } & \chi(\widehat{\tau}(s)(cv)) = 1 && \text{(definition } L_{\mathbb{S}}\text{)} \\ \text{iff } & \chi\big(\widehat{\tau}(\tau(s)(c))(v)\big) = 1 && \text{(definition } \widehat{\tau}\text{)} \\ \text{iff } & v \in L_{\mathbb{S}}(\tau(s)(c)) && \text{(definition } L_{\mathbb{S}}\text{)} \\ \text{iff } & v \in f(\tau(s)(c)) && \text{(induction hypothesis)} \\ \text{iff } & v \in \delta(f(s))(c) && (f \text{ is a morphism}) \\ \text{iff } & cv \in f(s) && \text{(definition } \delta\text{)} \end{aligned}$$

This suffices to prove the induction step of (3). QED

2.2 Properties of final coalgebras

Final coalgebras have various interesting properties. We first show that, if existing, final coalgebras are unique modulo isomorphism[3]. Because of this fact we will often speak of 'the' final T-coalgebra if T admits final coalgebras.

Proposition 2.4 *Let $\mathbb{Z} = (Z, \zeta)$ and $\mathbb{Z}' = (Z', \zeta')$ be final T-coalgebras for some functor $T : \mathsf{C} \to \mathsf{C}$. Then $\mathbb{Z}$ and $\mathbb{Z}'$ are isomorphic.*

Proof. By finality of $\mathbb{Z}$ there is a coalgebra morphism $g : \mathbb{Z}' \to \mathbb{Z}$, and by finality of $\mathbb{Z}'$ there is a coalgebra morphism $f : \mathbb{Z} \to \mathbb{Z}'$. But then the composition $g \circ f : \mathbb{Z} \to \mathbb{Z}$ is a coalgebra morphism as well, and by unicity it must be identical to the identity arrow $\mathsf{id}_{\mathbb{Z}}$. Similarly, we find that $g \circ f = \mathsf{id}_{\mathbb{Z}'}$. Thus $\mathbb{Z}$ and $\mathbb{Z}'$ are isomorphic indeed. QED

The following proposition states a key fact about final coalgebras.

Proposition 2.5 (Lambek's Lemma) *Let $\mathbb{Z}$ be a final T-coalgebra for some functor $T : \mathsf{C} \to \mathsf{C}$. Then the coalgebra map $\zeta : Z \to TZ$ of $\mathbb{Z}$ is an isomorphism in C.*

Proof. Applying the functor T to the coalgebra map ζ of $\mathbb{Z}$, we obtain the map $T\zeta : TZ \to TTZ$, and hence, a coalgebra $\mathbb{Z}_2 := (TZ, T\zeta)$. By finality of $\mathbb{Z}$ we obtain a coalgebra morphism $!$ from $\mathbb{Z}_2$ to $\mathbb{Z}$, given by a C-arrow $! : TZ \to Z$. But then the composition $! \circ \zeta$ is a coalgebra morphism from $\mathbb{Z}$ to itself, just like the identity arrow id_Z. In a diagram:

$$\begin{array}{ccccc} Z & \xrightarrow{\zeta} & TZ & \xrightarrow{!} & Z \\ {\scriptstyle\zeta}\downarrow & & \downarrow{\scriptstyle T\zeta} & & \downarrow{\scriptstyle\zeta} \\ TZ & \xrightarrow[T\zeta]{} & TTZ & \xrightarrow[T!]{} & TZ \end{array} \qquad \mathsf{id}_Z : Z \to Z \tag{4}$$

It follows by unicity that $! \circ \zeta = \mathsf{id}_Z$.

For the reverse composition $\zeta \circ !$ we have that $\zeta \circ ! = T! \circ T\zeta$ since $!$ is a morphism, cf. the right rectangle in the diagram above. But then we easily derive that $\zeta \circ ! = T(! \circ \zeta) = T\mathsf{id}_Z = \mathsf{id}_{TZ}$. In other words, $\zeta \circ !$ is the identity arrow on TZ.

Finally, since $! \circ \zeta = \mathsf{id}_Z$ and $\zeta \circ ! = \mathsf{id}_{TZ}$ we see that ζ is an isomorphism indeed, with $!$ as its inverse. QED

As an immediate corollary of this, we see that set functors involving the full powerset functor in a nontrivial way, will generally not admit a final coalgebra.

Corollary 2.6 *The categories of Kripke frames and of Kripke models do not admit final coalgebras.*

[3] See the appendix for the categorical definition of an isomorphism. In the case of coalgebras, two coalgebras $\mathbb{S}$ and $\mathbb{S}'$ are isomorphic if there are coalgebra morphisms $f : \mathbb{S} \to \mathbb{S}'$ and $g : \mathbb{S}' \to \mathbb{S}$ such that $g \circ f$ and $f \circ g$ are the identity arrows on $\mathbb{S}$ and $\mathbb{S}'$, respectively.

Proof. Recall that Kripke frames are coalgebras for the powerset functor P. Now suppose for contradiction that P admits a final coalgebra $\mathbb{Z} = (Z, \zeta)$. It would follow by Lambek's Lemma that ζ is a bijection between Z and its powerset; but this is impossible by Cantor's theorem. The case of Kripke models can be proved similarly. QED

The following proposition is one way to formalise the proof principle of *coinduction* — we shall come back to this.

Proposition 2.7 *Let $\mathbb{Z}$ be a final T-coalgebra for some set functor T. Then the relation $\simeq_{\mathbb{Z}}$ of behavioural equivalence on $\mathbb{Z}$ is the identity relation Δ_Z on Z:*

$$\simeq_{\mathbb{Z}} = \Delta_Z. \tag{5}$$

Proof. Suppose that z and z' are two states in $\mathbb{Z}$ that are behaviourally equivalent. In Remark 1.15 we saw that this implies that $\mathsf{beh}_{\mathbb{Z}}(z) = \mathsf{beh}_{\mathbb{Z}}(z')$, and since $\mathsf{beh}_{\mathbb{Z}}$ is the identity map on Z, it follows that $z = \mathsf{beh}_{\mathbb{Z}}(z) = \mathsf{beh}_{\mathbb{Z}}(z') = z'$. QED

2.3 Existence of final coalgebras

If not all set functors admit final coalgebras, which ones do? Some good sufficient conditions are known.

Definition 2.8 Let T be some set functor, and κ some cardinal. Call T *κ-small* if

$$T(S) = \bigcup\{(T\iota_S^A)[T(A)] \mid \iota_S^A : A \hookrightarrow S, |A| < \kappa\},$$

for all sets $S \neq \varnothing$, where for an arbitrary subset A of S, the arrow ι_S^A denotes the inclusion map of A into S. T is *small* if it is small for some cardinal κ. An ω-small functor is usually called *finitary*. $\lhd$

In words, the definition requires every element of $T(S)$ to be in the range of $T\iota$ for an appropriate inclusion map $\iota : A \hookrightarrow S$, where A is of size smaller than κ. In case T *preserves inclusions* (meaning that T maps any inclusion $\iota : A \hookrightarrow B$ to an inclusion $T\iota : TA \hookrightarrow TB$), the definition boils down to the requirement that

$$T(S) = \bigcup\{T(A) \mid A \subseteq S, |A| < \kappa\}.$$

Fact 2.9 *Every small set functor admits a final coalgebra.*

Examples of small functors abound; for instance, whenever we replace, in a Kripke polynomial functor, the power set functor by a bounded variant such as the finite power set functor, the result is a small functor.

In particular, the finite power set functor P_ω itself is ω-small. As an immediate corollary of this fact, the categories of image-finite frames and models, which can be represented as coalgebras for, respectively, the functors P_ω and $P\mathsf{Q} \times P_\omega$, both have final objects. More in general we can prove the following.

Corollary 2.10 *Every finitary Kripke polynomial functor admits a final coalgebra.*

For Set-based functors that do not admit a final coalgebra, there are various 'second-best' ways to proceed. For instance, one may show that T *does* have a final coalgebra in an *extended* or *modified* category.

Example 2.11 If one is willing to allow coalgebras taking a *class* rather than a set as their carrier, one may *create* a final coalgebra, outside the category Set, as follows. Let T be a set functor for which final coalgebras do not exist; for convenience we assume that all functors preserve inclusions.

Let SET be the category that has classes as objects, and class functions as arrows, that is, functions that map sets to sets but may have a class rather than a set as their (co-)domain. Call an endofunctor T on SET *set-based* if for each class C and each $X \in T\mathsf{C}$ there is a set $S \subseteq \mathsf{C}$ such that $X \in TS$. Now Aczel & Mendler proved that every set-based endofunctor on SET admits a final coalgebra – the similarity to Fact 2.9 is no coincidence.

This fact can be used as follows. Given an endofunctor T on Set, there is a *unique* way to extend T to a set-based endofunctor T^+ on SET. (On objects, simply put $T^+(\mathsf{C}) := \bigcup\{T(S) \mid S \subseteq \mathsf{C} \text{ a set}\}$.)

The theorem of Aczel & Mendler then guarantees the existence of a final object $\mathbb{Z}$ in $\mathsf{Coalg}(T^+)$. This coalgebra will be class-based if T does not admit a final coalgebra, but it will be final, not only with respect to the set-based coalgebras in $\mathsf{Coalg}(T^+)$, but also with respect to the class-based ones. As an important manifestation of this idea, Aczel showed that the class of non-well-founded sets provides the final coalgebra for (the SET-based extension of) the power set functor.

Example 2.12 One way to look at Lambek's Lemma is that final T-coalgebras provide solutions to the 'equation' $S \cong TS$. In the case of the powerset functor, Cantor's theorem states that this equation does not have a solution in the category Set. This situation is reminiscent of that in *domain theory*, which provides solutions to the equation $X \cong X^X$ by imposing topological structure on sets.

Something similar can be done here. Define a *Stone space* to be a pair $\mathbb{X} = (X, \tau)$, where τ is a zero-dimensional compact Hausdorff space, and let Stone denote the category of Stone spaces as objects with continuous maps as arrows. As an analog to the powerset functor on Set, we can define the *Vietoris* functor V on the category Stone; on objects, the Vietoris space $V\mathbb{X}$ is based on the collection of *compact* subsets of X. We may then show that, indeed, the final coalgebra for this functor exists.

Further on we will see that in modal logic, whereas the canonical model (over a finite set Q of proposition letters) is not final in the category of coalgebras for the Kripke model functor $K_{P\mathsf{Q}} \times P$, we may identify the canonical *general* model over Q with the final $K_{P\mathsf{Q}} \times V$-coalgebra. Thus we can solve the equation $S \cong TS$ by *modifying* the base category of our coalgebras.

2.4 The terminal sequence

Whether the functor admits a final coalgebra or not, one may always (try to) approximate it by considering the so-called final or terminal *sequence*.

Example 2.13 Let us first consider an example outside the category Set. Suppose that the poset $\mathbb{C} = (C, \leq)$ is in fact a complete lattice, that is, with each subset $X \subseteq C$ we may associate a *meet* or *greatest lower bound* $\bigwedge X$; this means in particular that $\mathbb{C}$ is bounded: it has a largest element $\top := \bigwedge \varnothing$, and a smallest element $\bot := \bigvee \varnothing$.

The Knaster-Tarski theorem states that in this setting, every monotone map $F : C \to C$ has both a least and a greatest fixpoint. That is, every endofunctor F on the *category* $\mathbb{C}$ admits both an initial and a final coalgebra (see Example 2.1).

It is instructive for our purposes to prove this theorem, and in particular, to see how to find the greatest fixpoint by approximating it from above. We define an ordinal-indexed sequence $\langle z_\alpha \rangle$ using transfinite induction:

$$z_\alpha := \begin{cases} \top & \text{if } \alpha = 0 \\ Fz_\beta & \text{if } \alpha = \beta + 1 \text{ is a successor ordinal} \\ \bigwedge\{z_\beta \mid \beta < \alpha\} & \text{if } \alpha \text{ is a limit ordinal} \end{cases}$$

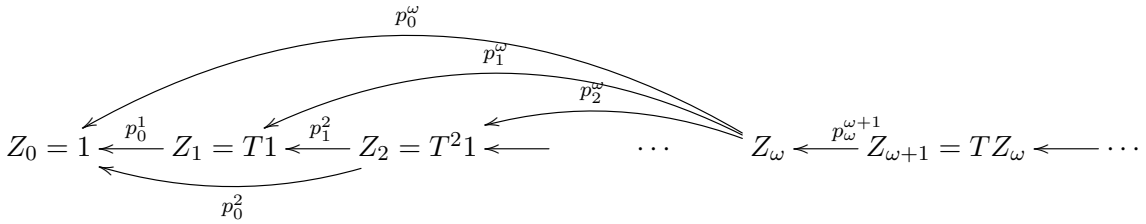
Note that in fact, if we take 0 to be a limit ordinal, we can reduce the first clause of the definition to the third.

It is not hard to prove, for a monotone map $F : C \to C$, the existence of an ordinal α for which $z_\alpha = z_{\alpha+1}$, and to show that the object z_α is in fact the greatest fixpoint of F.

In the case of set functors, we may take inspiration from this example to define the *terminal sequence* associated with T.

Definition 2.14 The *final* or *terminal sequence* associated with a given set functor T, is an ordinal indexed sequence of objects $\langle Z_\alpha \rangle$ with maps $p^\alpha_\beta : Z_\alpha \to Z_\beta$ for $\beta \leq \alpha$, such that (i) $Z_{\alpha+1} = TZ_\alpha$ and $p^{\alpha+1}_{\beta+1} = Tp^\alpha_\beta$, (ii) $p^\alpha_\alpha = \mathsf{id}_{Z_\alpha}$ and $p^\beta_\gamma \circ p^\alpha_\beta = p^\alpha_\gamma$, (iii) if λ is a limit ordinal, then Z_λ with $\{p^\lambda_\alpha \mid \alpha < \lambda\}$ is a limit[4] of the diagram with objects $\{Z_\alpha \mid \alpha < \lambda\}$ and arrows $\{p^\alpha_\beta \mid \alpha, \beta < \lambda\}$. (In particular, taking 0 to be a limit ordinal, we find that $Z_0 = 1$ is some final object of the category Set, i.e., a singleton set.) $\lhd$

In the diagram below we depict an initial part of this construction:



(6)

It is not hard to prove that, modulo isomorphism, the terminal sequence is uniquely determined by these conditions. Intuitively, it can be seen as an approximation of the final coalgebra for T. That is, where elements of the final coalgebra represent 'complete' behavior, elements of Z_α represent behavior that can be 'performed in α many steps'.

[4]See the appendix for the categorical definition of a limit.

To make this precise and formal, observe that for any T-coalgebra $\mathbb{S}$ there is a *unique* ordinal-indexed class of functions $\mathsf{beh}_\alpha : S \to Z_\alpha$ such that beh_0 is fixed by the finality of Z_0 in Set, $\mathsf{beh}_{\alpha+1} := (T\mathsf{beh}_\alpha) \circ \sigma$:

$$\cdots \qquad \begin{array}{ccc} Z_\alpha & \xleftarrow{p_\alpha^{\alpha+1}} & Z_{\alpha+1} \\ {\scriptstyle \mathsf{beh}_\alpha}\uparrow & \nearrow_{\mathsf{beh}_{\alpha+1}} & \uparrow{\scriptstyle T\mathsf{beh}_\alpha} \\ S & \xrightarrow[\sigma]{} & TS \end{array} \qquad \cdots \tag{7}$$

while for a limit ordinal λ, beh_λ is given as the unique map $\mathsf{beh}_\lambda : S \to Z_\lambda$ such that $\mathsf{beh}_\alpha = p_\alpha^\lambda \circ \mathsf{beh}_\lambda$ for all $\alpha < \lambda$. It is not hard to prove that, for instance, $\mathbb{S}, s \simeq \mathbb{S}', s'$ implies that $\mathsf{beh}_\alpha(s) = \mathsf{beh}_\alpha(s')$ for all α.

The relation with final coalgebras can be made precise, as follows. On the one hand, if the terminal sequence converges, in the sense that some arrow $p_\alpha^{\alpha+1}$ is a bijection, then the coalgebra $(Z_\alpha, (p_\alpha^{\alpha+1})^{-1})$ is a final coalgebra for T. And conversely, under some mild constraints on T, Adámek & Koubek proved that if T admits a final coalgebra, then the terminal sequence converges to it.

2.5 Coinduction as a definition principle

Coinduction is an important coalgebraic principle, and just like its algebraic counterpart of *induction*, it can be used as a tool to *define* various operations, but also as a coalgebraic *proof* principle.

To see how coinductive definitions work, suppose that $\mathbb{Z} = (Z, \zeta)$ is the final coalgebra for some set functor T. Coinduction is based on the observation that, in order to define a map from some set S to Z, it suffices to turn S into a coalgebra by endowing it with coalgebra structure: any coalgebra map $\sigma : S \to TS$ canonically induces a map from S to Z, namely the unique coalgebra morphism $!_\sigma : (S, \sigma) \to \mathbb{Z}$. (In fact, we saw this principle already at work in the proof of Lambek's Lemma, where we defined a map from TZ to Z by considering the coalgebra map $T\zeta$ on TZ.)

Example 2.15 Let $\mathbb{Z} = (C^\omega, \mathsf{h}, \mathsf{t})$ be the *stream coalgebra* of Example 1.10 — here we write h and t rather than head and tail. We already saw that $\mathbb{Z}$ is the final $K_C \times Id$-coalgebra, we can now use the finality to define operations on streams.

To start with, consider the coalgebra map $\langle \mathsf{h}, \mathsf{t} \circ \mathsf{t} \rangle : C^\omega \to C \times C^\omega$ defined as $\gamma_e(\alpha) := (\mathsf{h}(\alpha), \mathsf{t}(\mathsf{t}(\alpha)))$. By finality of $\mathbb{Z}$ there is a unique map $\mathsf{e} : C^\omega \to C^\omega$ making the following diagram commute:

$$\begin{array}{ccc} C^\omega & \xrightarrow{\mathsf{e}} & C^\omega \\ {\scriptstyle \langle \mathsf{h}, \mathsf{t}\circ\mathsf{t} \rangle}\downarrow & & \downarrow{\scriptstyle \langle \mathsf{h}, \mathsf{t} \rangle} \\ C \times C^\omega & \xrightarrow[\mathsf{id}_C \times \mathsf{e}]{} & C \times C^\omega \end{array} \tag{8}$$

Another way of looking at this definition is that, in order to define $\mathsf{e}(\alpha)$, we specfify

$$\begin{array}{lcl} \mathsf{h}(\mathsf{e}(\alpha)) & := & \mathsf{h}(\alpha) \\ \mathsf{t}(\mathsf{e}(\alpha)) & := & \mathsf{e}(\mathsf{t}(\mathsf{t}(\alpha))). \end{array}$$

In fact, the map $\mathsf{e} : C^\omega \to C^\omega$ is the operation on streams that creates a new stream out of all items at an *even* position in the input stream. To prove this, it suffices to show that the map $\mathsf{e} := \lambda\alpha.(\lambda n.\alpha(2n))$ makes the diagram (8) commute.

Similary, we can define a map $\mathsf{q} : C^\omega \to C^\omega$ selecting the *odd* items of an input stream, by means of the following diagram:

$$\begin{array}{ccc} C^\omega & \xrightarrow{\mathsf{q}} & C^\omega \\ {\scriptstyle \langle \mathsf{h}\circ\mathsf{t}, \mathsf{t}\circ\mathsf{t}\rangle}\downarrow & & \downarrow{\scriptstyle \langle \mathsf{h},\mathsf{t}\rangle} \\ C \times C^\omega & \xrightarrow[\mathsf{id}_C \times \mathsf{q}]{} & C \times C^\omega \end{array} \tag{9}$$

Example 2.16 Fix an alphabet C. Recall from Proposition 2.3 that the language coalgebra $\mathbb{L} = (\mathcal{L}, \omega, \delta)$ is the final coalgebra for the 'automaton' functor $2 \times Id^C$. We can use this fact to define operations on languages.

For instance, given a word $u = c_1 \cdots c_k$ (with $k \geq 0$), we let $\otimes u$ denote its converse, $\otimes u := c_k \cdots c_1$, and we set $\otimes L := \{\otimes u \mid u \in L\}$. Coinductively, we can define this language by imposing the following structure on $\mathcal{L}$. As the acceptance condition we simply take the same map ω as for $\mathbb{L}$, while for the transition map τ we put

$$\tau(L)(c) := \{u \in C^* \mid uc \in L\}.$$

We leave it as an exercise for the reader to verify that the following diagram commutes:

$$\begin{array}{ccc} \mathcal{L} & \xrightarrow{\otimes} & \mathcal{L} \\ {\scriptstyle \langle \omega, \tau\rangle}\downarrow & & \downarrow{\scriptstyle \langle \omega, \delta\rangle} \\ 2 \times \mathcal{L}^C & \xrightarrow[\mathsf{id}_2 \times \otimes^C]{} & 2 \times \mathcal{L}^C \end{array} \tag{10}$$

so that by finality, $\otimes : \mathcal{L} \to \mathcal{L}$ is the *unique* coalgebra morphism $\otimes : (\mathcal{L}, \tau, \omega) \to \mathbb{L}$. That means that (10) can be seen as a coinductive *definition* of $\otimes$.

Example 2.17 As a slightly different example, we give a coinductive definition of the *shuffle* $K \,||\, L$ of two languages K and L. For an intuitive description of this operation[5], a word belongs to $K \,||\, L$ if it consists of two interleaved words, from K and L, respectively. In particular, we want that

$$(K \,||\, L)_c = K_c \,||\, L \;\cup\; K \,||\, L_c,$$

[5]For a more precise definition of the language $K \,||\, L$, first define the relation $\lhd$ on finite words by putting $c_1 \cdots c_k \lhd d_1 \cdots d_n$ if there is an order-preserving map on the indices $f : \{1, \ldots, k\} \to \{1, \ldots, n\}$ such that $c_i = d_{fi}$, for all $i \in \{1, \ldots, k\}$. Say that w *merges* u and v if $|w| = |u| + |v|$ and both $u \lhd w$ and $v \lhd w$. We can then define $K \,||\, L$ as the collection of all words that merge words from K and L.

for all $c \in C$, where we recall that, for a language $L \in \mathcal{L}$, we let L_a denote its a-derivative, $L_a := \{u \in C^* \mid au \in L\}$.

For a coinductive definition, we consider the set $\mathcal{E}$ of expressions defined by the following grammar:

$$E \ ::= \ \underline{L} \mid E_0 + E_1 \mid E_0 \star E_1$$

where $L \in \mathcal{L}$, i.e., we associate a formal symbol $\underline{L}$ with every language L.

To turn the set $\mathcal{E}$ into an automaton $\mathbb{E} := (\mathcal{E}, \chi, \tau)$, consider the following axioms and rules, which inductively define the relation $\downarrow$ and the functions $\xrightarrow{c}$, with $c \in C$. Here we use the notation introduced in the beginning of section 2.1.

$$\frac{}{\underline{L}\downarrow}\,\varepsilon \in L \qquad \frac{E\downarrow}{(E+F)\downarrow} \qquad \frac{F\downarrow}{(E+F)\downarrow} \qquad \frac{E\downarrow \quad F\downarrow}{(E\star F)\downarrow}$$

$$\frac{}{\underline{L} \xrightarrow{c} \underline{L_c}} \qquad \frac{E \xrightarrow{c} E' \quad F \xrightarrow{c} F'}{E + F \xrightarrow{c} E' + F'} \qquad \frac{E \xrightarrow{c} E' \quad F \xrightarrow{c} F'}{E \star F \xrightarrow{c} E' \star F + E \star F'}$$

It is not hard to see that this deductive system uniquely determines two operations, $\chi : \mathcal{E} \to 2$ and $\tau : \mathcal{E} \to \mathcal{E}^C$. Hence by finality of $\mathbb{L}$ there is a unique coalgebra morphism $f : \mathbb{E} \to \mathbb{L}$. Think of $f(E)$ as the *interpretation* of the term E. We can then define

$$K \,||\, L := f(\underline{K} \star \underline{L}).$$

To get a feeling for this definition we compute the derivative $(K \,||\, L)_c$:

$$\begin{aligned}
(K \,||\, L)_c &= f(\underline{K} \star \underline{L})_c && \text{(definition } ||\text{)} \\
&= \delta(f(\underline{K} \star \underline{L}))(c) && \text{(definition } \delta\text{)} \\
&= f(\tau(\underline{K} \star \underline{L})(c)) && (f \text{ is a morphism}) \\
&= f(\underline{K_c} \star \underline{L} + \underline{K} \star \underline{L_c}) && \text{(definition } \tau\text{)} \\
&= f(\underline{K_c} \star \underline{L}) \cup f(\underline{K} \star \underline{L_c}) && (*) \\
&= K_c \,||\, L \cup K \,||\, L_c && \text{(definition } ||\text{)}
\end{aligned}$$

Here we use in (*) the observation that $f(E + F) = f(E) \cup f(F)$, which is easily proved by coinduction, see Example 2.20.

2.6 Coinduction as a proof principle

As a *proof* principle, coinduction has two manifestations. In its most direct form, proofs by coinduction use the uniqueness of coalgebra morphisms to a final coalgebra.

Example 2.18 Consider the maps e and q of Example 2.15. We claim that

$$\mathsf{q} = \mathsf{e} \circ \mathsf{t}, \tag{11}$$

as should be clear intuitively. To prove (11) coinductively, it suffices to prove that the map $\mathsf{e} \circ \mathsf{t} : C^\omega \to C^\omega$, just like q, is a coalgebra morphism $\mathsf{e} \circ \mathsf{t} : (C^\omega, \langle \mathsf{h} \circ \mathsf{t}, \mathsf{t} \circ \mathsf{t} \rangle) \to (C^\omega, \langle \mathsf{h}, \mathsf{t} \rangle)$; that is, the diagram below commutes:

$$\begin{array}{ccc} C^\omega & \xrightarrow{\mathsf{e}\circ\mathsf{t}} & C^\omega \\ {\scriptstyle \langle \mathsf{h}\circ\mathsf{t}, \mathsf{t}\circ\mathsf{t}\rangle}\downarrow & & \downarrow{\scriptstyle \langle \mathsf{h},\mathsf{t}\rangle} \\ C \times C^\omega & \xrightarrow[\mathsf{id}_C \times \mathsf{e}\circ\mathsf{t}]{} & C \times C^\omega \end{array} \tag{12}$$

But for this purpose it suffices to show that the following two equations hold:

$$\mathsf{h} \circ (\mathsf{e} \circ \mathsf{t}) = \mathsf{h} \circ \mathsf{t} \tag{13}$$

$$\mathsf{t} \circ (\mathsf{e} \circ \mathsf{t}) = (\mathsf{e} \circ \mathsf{t}) \circ (\mathsf{t} \circ \mathsf{t}) \tag{14}$$

This is not so hard. For (13), we may calculate

$$\begin{aligned} \mathsf{h} \circ (\mathsf{e} \circ \mathsf{t}) &= (\mathsf{h} \circ \mathsf{e}) \circ \mathsf{t} && \text{(associativity)} \\ &= (\mathsf{id}_C \circ \mathsf{h}) \circ \mathsf{t} && \text{(diagram (8))} \\ &= \mathsf{h} \circ \mathsf{t}, && (\mathsf{id}_C \circ \mathsf{h} = \mathsf{h}) \end{aligned}$$

while we prove (14) as follows:

$$\begin{aligned} \mathsf{t} \circ (\mathsf{e} \circ \mathsf{t}) &= (\mathsf{t} \circ \mathsf{e}) \circ \mathsf{t} && \text{(associativity)} \\ &= (\mathsf{e} \circ (\mathsf{t} \circ \mathsf{t})) \circ \mathsf{t} && \text{(diagram (8))} \\ &= (\mathsf{e} \circ \mathsf{t}) \circ (\mathsf{t} \circ \mathsf{t}) && \text{(associativity)} \end{aligned}$$

Example 2.19 In a similar way we can prove that

$$\mathsf{zip} \circ \langle \mathsf{e}, \mathsf{q} \rangle = \mathsf{id}_{C^\omega}, \tag{15}$$

where zip is the map defined in Example 1.12.

By finality it suffices to show that $\mathsf{zip} \circ \langle \mathsf{e}, \mathsf{q} \rangle$ is a coalgebra morphism on the stream coalgebra, and since we know that zip is a coalgebra morphism (i.e., the right rectangle below commutes), we can confine ourselves to proving that the left rectangle in the diagram below commutes:

$$\begin{array}{ccccc} C^\omega & \xrightarrow{\langle \mathsf{e},\mathsf{q}\rangle} & C^\omega \times C^\omega & \xrightarrow{\mathsf{zip}} & C^\omega \\ {\scriptstyle \langle \mathsf{h},\mathsf{t}\rangle}\downarrow & & \downarrow{\scriptstyle \delta} & & \downarrow{\scriptstyle \langle \mathsf{h},\mathsf{t}\rangle} \\ C \times C^\omega & \xrightarrow[T\langle \mathsf{e},\mathsf{q}\rangle]{} & C \times (C^\omega \times C^\omega) & \xrightarrow[T\mathsf{zip}]{} & C \times C^\omega \end{array}$$

Here $T\langle \mathsf{e}, \mathsf{q} \rangle = \mathsf{id}_C \times \langle \mathsf{e}, \mathsf{q} \rangle$, $T\mathsf{zip} = \mathsf{id}_C \times \mathsf{zip}$, and δ is as given in Example 1.12: $\delta(\alpha, \beta) := (\mathsf{h}(\alpha), (\beta, \mathsf{t}(\alpha))$.

To verify that the left rectangle above commutes we need to check that $\delta \circ \langle \mathsf{e}, \mathsf{q} \rangle = T\langle \mathsf{e}, \mathsf{q} \rangle \circ \langle \mathsf{h}, \mathsf{t} \rangle$, which boils down to proving

(a) $\mathsf{h} \circ \mathsf{e} = \mathsf{h}$,

(b) $\mathsf{q} = \mathsf{e} \circ \mathsf{t}$
(c) $\mathsf{t} \circ \mathsf{e} = \mathsf{q} \circ \mathsf{t}$.
But we obtain (a) because $\mathsf{h} \circ \mathsf{e} = \mathsf{id}_C \circ \mathsf{h}$ (definition of e), and (b) was shown in the previous example, cf. (11). Finally, for (c), observe that $\mathsf{t} \circ \mathsf{e} = \mathsf{e} \circ (\mathsf{t} \circ \mathsf{t})$ by definition of e (diagram (8)), and $\mathsf{e} \circ (\mathsf{t} \circ \mathsf{t}) = \mathsf{q} \circ \mathsf{t}$ by associativity and (11).

Often, coinduction is referred to as the proof principle that uses the fact that behavioural equivalence is the identity relation on a final coalgebra (Proposition 2.7). More specifically, given the fact that bisimilarity implies behavioural equivalence (as we will see further on), one may prove two states in a final coalgebra to be identical if we can link them by a bisimulation.

Example 2.20 In the case of C-automata (coalgebras of type $2 \times Id^C$), a bisimulation on a coalgebra (S, χ, τ) is a relation $B \subseteq S \times S$ such that, whenever $(s_0, s_1) \in B$, we have
(acc) $\chi(s_0) = \chi(s_1)$ (that is: $s_0{\downarrow}$ iff $s_1{\downarrow}$), and
(nxt) $(\tau(s_0)(c), \tau(s_1)(c)) \in B$, for all $c \in C$.

Let us now prove the statement that (cf. Example 2.17)

$$f(\underline{L}) = L, \tag{16}$$

for every language $L \in \mathcal{L}$. By coinduction, it suffices to show that the relation

$$B := \{(f(\underline{L}), L) \mid L \in \mathcal{L}\} \tag{17}$$

is a bisimulation. We check the two conditions.

For (acc), we observe that $f(\underline{L}){\downarrow}$ (in $\mathbb{L}$) iff $\underline{L}{\downarrow}$ (in $\mathbb{E}$) since f is a coalgebra morphism. But we have $\underline{L}{\downarrow}$ in $\mathbb{E}$ iff $\varepsilon \in L$ by definition of acceptance in $\mathbb{E}$, and we have $L{\downarrow}$ (in $\mathbb{L}$) iff $\varepsilon \in L$ by definition of acceptance in $\mathbb{L}$. This suffices to prove (acc).

For (nxt) we need to show, for an arbitrary language $L \in \mathcal{L}$ and an arbitrary letter $c \in C$, that the pair $\big(\delta(f(\underline{L}), c), \delta(L, c)\big) \in B$. To that aim, observe that $\delta(f(\underline{L}), c) = f(\tau(\underline{L})(c)) = f(\underline{L_c})$, respectively since f is a morphism and by definition of τ. But since we have $\delta(L)(c) = L_c$ (by definition of δ), it is immediate that $\big(\delta(f(\underline{L}), c), \delta(L, c)\big) = \big(f(\underline{L_c}), L_c\big) \in B$. This finishes the proofs of (17) and (16).

Similarly, we can prove that

$$f(E_0 + E_1) = f(E_0) \cup f(E_1), \tag{18}$$

for all expressions E_0 and E_1, by showing that the relation

$$R := \Big\{ \big(f(E_0 + E_1), f(E_0) \cup f(E_1)\big) \mid E_0, E_1 \in \mathcal{E} \Big\}$$

is a bisimulation on $\mathbb{L}$.

Finally, we can use coinduction to establish, in a relatively straightforward way, various useful properties of the operation $||$, such as commutativity, associativity, or distribution with respect to $+/\cup$.

3 Bisimilarity and Behavioural Equivalence

In section 1.4 of the Introduction we defined two coalgebraic notions of equivalence: behavioral equivalence and bisimilarity. In this chapter we discuss these notions in more detail.

3.1 Basic observations

Obviously, the first question is how the notions of behavioral equivalence and bisimilarity relate to each other. One direction is clear: bisimilarity is a sufficient condition for behavioral equivalence.

Proposition 3.1 *Let $T : \mathsf{Set} \to \mathsf{Set}$ be some functor, and let s_0 and s_1 be states of the T-coalgebras $\mathbb{S}_0$ and $\mathbb{S}_1$, respectively. Then $\mathbb{S}_0, s_0 \leftrightarroweq \mathbb{S}_1, s_1$ implies $\mathbb{S}_0, s_0 \simeq \mathbb{S}_1, s_1$.*

Proof. In the special case that T admits a final coalgebra, a very simple proof obtains. Assume that $\mathbb{S}_0, s_0 \leftrightarroweq \mathbb{S}_1, s_1$, and let $B \subseteq S_0 \times S_1$ with $\beta : B \to TB$ be a coalgebra witnessing this. It follows from the definitions that both $\mathsf{beh}_{\mathbb{S}_0} \circ \pi_0$ and $\mathsf{beh}_{\mathbb{S}_1} \circ \pi_1$ are coalgebraic morphisms from (B, β) to the final coalgebra, so from finality it follows that $\mathsf{beh}_{\mathbb{S}_0} \circ \pi_0 = \mathsf{beh}_{\mathbb{S}_1} \circ \pi_1$. From this it is immediate that $B \subseteq \simeq$; and so from $(s_0, s_1) \in B$ it follows that $\mathbb{S}_0, s_0 \simeq \mathbb{S}_1, s_1$.

In the general case the proof of this proposition is similar to the one of Theorem 3.15 below (with an application of *pushouts* instead of pullbacks), so we omit details. QED

The converse statement of Proposition 3.1 does not hold: in general, bisimilarity is a strictly stronger notion than behavioral equivalence.

Example 3.2 Consider the so-called '3-2-functor' $T_2^3 : \mathsf{Set} \to \mathsf{Set}$ given on objects by

$$T_2^3(S) := \big\{(s_0, s_1, s_2) \in S^3 \mid |\{s_0, s_1, s_2\}| \leq 2\},$$

while for an arrow $f : S \to S'$ we define $(T_2^3 f)(s_0, s_1, s_2) := (fs_0, fs_1, fs_2)$. We leave it as an exercise to the reader to verify that this indeed defines a set functor.

Now consider the following coalgebra $\mathbb{S} = (S, \sigma)$, where $S = \{0, 1\}$ and σ is given by $\sigma(0) = (0, 0, 1)$ and $\sigma(1) = (1, 0, 0)$. Then it is not hard to see that $\mathbb{S}, 0 \simeq \mathbb{S}, 1$, but at the same time we claim that there is *no* T_2^3-bisimulation on $\mathbb{S}$ linking 0 and 1. To see this, suppose for contradiction that $R \subseteq S \times S$ would be such a bisimulation, witnessed by the coalgebra map $\rho : R \to T_2^3 R$. If the projection maps $\pi_0, \pi_1 : R \to S$ are to be coalgebra morphisms, ρ has to map the pair $(0, 1)$ to some triple $\rho(0, 1) = \big((s_0, s_1), (t_0, t_1), (u_0, u_1)\big)$ such that

$$\begin{array}{rcccccccc} (s_0, t_0, u_0) & = & ((T_2^3\pi_0) \circ \rho)(0,1) & = & (\sigma \circ \pi_0)(0,1) & = & \sigma(0) & = & (0,0,1) \\ (s_1, t_1, u_1) & = & ((T_2^3\pi_1) \circ \rho)(0,1) & = & (\sigma \circ \pi_1)(0,1) & = & \sigma(1) & = & (1,0,0). \end{array}$$

Clearly then we find $\rho(0, 1) = \big((s_0, s_1), (t_0, t_1), (u_0, u_1)\big) = \big((0, 1), (0, 0), (1, 0)\big)$. But this object does not belong to the set $T_2^3 R$, since (s_0, s_1), (t_0, t_1) and (u_0, u_1) are all distinct.

Example 3.3 A more natural example of a set functor for which behavioural equivalence and bisimilarity are properly distinct notions, is the *monotone neighbourhood functor* M. We will come back to this example later.

In section 3.3 below we will discuss an important class of set functors for which we do have $\simeq = \leftrightarroweq$. First, however, we make some basic observations on bisimulations and, in the next section, we give an alternative characterization of bisimulations.

Example 3.4 For an arbitrary set functor T, it is easy to see that for any coalgebra $\mathbb{S}$, the diagonal relation Δ_S is a bisimulation equivalence on $\mathbb{S}$. Furthermore, the converse of a bisimulation is again a bisimulation.

As another general example, coalgebra morphisms can be seen as functional bisimulations. To be more precise, let $f : S_0 \to S_1$ be a function between the carriers of two T-coalgebras $\mathbb{S}_0$ and $\mathbb{S}_1$. Recall that the *graph* of f is the relation $\mathsf{Gr}f := \{(s, f(s)) \mid s \in S_0\}$. Then it holds that

$$f \text{ is a coalgebra morphism iff its graph } \mathsf{Gr}f \text{ is a bisimulation.} \tag{19}$$

In order to see why this is so, first suppose that $\mathsf{Gr}f : \mathbb{S}_0 \leftrightarroweq \mathbb{S}_1$. Since the projection map $\pi_0 : \mathsf{Gr}f \to S_0$ is a bijective morphism, its inverse π_0^{-1} is also a morphism. But then $f = \pi_1 \circ \pi_0^{-1}$, as the composition of two morphisms, is also a morphism. For the other direction, suppose that f is a morphism; then it is straightforward to verify that the map $(T\pi_0)^{-1} \circ \sigma \circ \pi_0$ equips the set $\mathsf{Gr}f$ with the required coalgebraic structure.

However, the collection of bisimulations is not in general closed under taking relational composition, and the relation $\leftrightarroweq$ of bisimilarity on a given coalgebra is generally not an equivalence relation.

Remark 3.5 Recall that in Remark 1.20 we define a relation $Z \subseteq S_0 \times S_1$ to be a bisimulation between two *T-models* (S_0, σ_0, V_0) and (S_1, σ_1, V_1) over some set Q of proposition letters, if it is a standard bisimulation between the two structures, seen as coalgebras for the functor $T_{\mathsf{Q}} = K_{P\mathsf{Q}} \times T$.

We leave it for the reader to verify that this is equivalent to requiring that such a relation Z is a bisimulation between the T-coalgebras (S_0, σ_0) and (S_1, σ_1) which meets (as in Example 1.3) the condition (atom) that s and s' satisfy the same proposition letters in Q, when linked by Z.

3.2 Bisimulations and relation lifting

Bisimulations admit an elegant alternative characterization which involves the notion of *relation lifting*.

Example 3.6 As an example, consider the power set functor P. Recall that a relation $B \subseteq S_0 \times S_1$ is a bisimulation between two P-coalgebras (Kripke frames) $\mathbb{S}_0 = (S_0, R_0[\cdot])$ and $\mathbb{S}_1 = (S_1, R_1[\cdot])$ iff B satisfies the conditions *(back)* and *(forth)* of Example 1.3. Now suppose that we define, for an arbitrary relation $R \subseteq S_0 \times S_1$, the relation $\overline{P}(R) \subseteq P(S_0) \times P(S_1)$ by putting

$$\overline{P}(R) := \{(Q_0, Q_1) \mid \forall q_0 \in Q_0\, \exists q_1 \in Q_1.\, (q_0, q_1) \in R \text{ and } \forall q_1 \in Q_1\, \exists q_0 \in Q_0.\, (q_0, q_1) \in R\}. \tag{20}$$

In other words, we *lift* the relation R to the level of the power sets of S_0 and S_1. The definition of a bisimulation between P-coalgebras can now be characterized as follows:

$$B : \mathbb{S}_0 \leftrightarrows \mathbb{S}_1 \text{ iff } (R_0[s_0], R_1[s_1]) \in \overline{P}(B) \text{ for all } (s_0, s_1) \in B.$$

This nice way of characterizing bisimulation via relation lifting is not limited to the power set functor — it applies in fact to *every* set functor.

Definition 3.7 Let T be some set functor. Given a relation $R \subseteq S_0 \times S_1$, consider R as a *span*

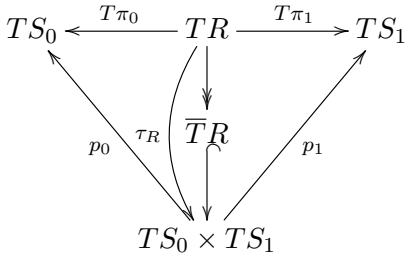
$$S_0 \xleftarrow{\pi_0} R \xrightarrow{\pi_1} S_1 \ ,$$

where $\pi_i : R \to S_i$ and $p_i : TS_0 \times TS_1 \to TS_i$ denote the respective projection maps. We define the *relation lifting* of R as the relation $\overline{T}R \subseteq TS_0 \times TS_1$ given by

$$\overline{T}R := \{((T\pi_0)(u), (T\pi_1)(u)) \mid u \in TR\}, \tag{21}$$

that is, $\overline{T}R$ is the image of TR under the map $\tau_R := \langle T\pi_0, T\pi_1 \rangle$. $\lhd$

In other words, we apply the functor T to the relation R, seen as a span. It follows from the category-theoretic properties of the product $TS_0 \times TS_1$ that there is a unique map $\tau_R := \langle T\pi_0, T\pi_1 \rangle$ from TR to $TS_0 \times TS_1$ such that $p_i \circ \tau_R = T\pi_i$ for $i = 0, 1$. Now we define $\overline{T}R$ as the image of TR under the map τ_R obtained from the lifted projection maps $T\pi_0$ and $T\pi_1$. In a diagram:



Remark 3.8 Strictly speaking, the definition of the relation lifting of a given relation depends on how we type it. That is, given sets S_0, S_1, X_0 and X_1, and a relation R such that $R \subseteq S_0 \times S_1$ and $R \subseteq X_0 \times X_1$, for the definition of TR it matters whether we look at R as a relation from S_0 to S_1 or as a relation from X_0 to X_1.

This possible source of ambiguity evaporates if we require the functor T to be standard, that is: maps inclusions to inclusions; see the appendix for more details. In this chapter we will only use the lifting of a relation in a setting where its type is fixed, but readers who are worried about this issue may add the (fairly harmless) condition that T is standard.

The results listed in the following theorems summarize the most important properties of bisimulations.

Theorem 3.9 *Let $\mathbb{S}_0$ and $\mathbb{S}_1$ be two coalgebras for some set functor T. Then the following holds, for any relation $B \subseteq S_0 \times S_1$:*

$$B : \mathbb{S}_0 \leftrightarrows \mathbb{S}_1 \textit{ iff } (\sigma_0(s_0), \sigma_1(s_1)) \in \overline{T}(B), \textit{ for all } (s_0, s_1) \in B. \tag{22}$$

Proof. The proof of this theorem is an almost immediate consequence of the definitions. The key observation is that both statements in (22) are equivalent to the requirement that for every pair $(s_0, s_1) \in B$ there is a $u \in TB$ such that $T\pi_i(u) = \sigma_i(s_i)$, for each $i = 0, 1$. QED

Theorem 3.10 *Let $\mathbb{S}_0$ and $\mathbb{S}_1$ be two coalgebras for some set functor T. Then*

1. *the collection of bisimulations between $\mathbb{S}_0$ and $\mathbb{S}_1$ forms a complete lattice under the inclusion order, with joins given by unions;*
2. *the bisimilarity relation $\leftrightarrows$ is the largest bisimulation between $\mathbb{S}_0$ and $\mathbb{S}_1$.*

Proof. The crucial observation in the proof of this theorem is that

$$\overline{T} : P(S_0 \times S_1) \to P(TS_0 \times TS_1) \text{ is a monotone operation.} \tag{23}$$

For a proof, let $R \subseteq R'$ be two relations between S_0 and S_1, with $\iota : R \to R'$ denoting the inclusion map. By definition of $\overline{T}$, we may without loss of generality represent an arbitrary element of $\overline{T}(R)$ as a pair $\tau_R(u) = ((T\pi_0)(u), (T\pi_1)(u))$ for some $u \in TR$. Define $u' := (T\iota)(u)$, then u' belongs to TR', and for each i we find that $(T\pi'_i)(u') = (T\pi'_i \circ T\iota)(u) = T(\pi'_i \circ \iota)(u) = (T\pi_i)(u)$. That is, $\tau_R(u) = \tau_{R'}(u')$, which shows that $\tau_R(u)$ belongs to $\overline{T}R'$. This proves (23).

Now for the proof of part 1, recall that a partial order is a complete lattice if it is closed under arbitrary joins. Hence, it suffices to prove that the union B of a collection $\{B_j \mid j \in J\}$ of bisimulations is again a bisimulation. Take an arbitrary pair $(s_0, s_1) \in B$. Then (s_0, s_1) belongs to B_j for some $j \in J$. Hence, by Theorem 3.9, we find $(\sigma_0 s_0, \sigma_1 s_1)$ in $\overline{T}(B_j)$, so $(\sigma_0(s_0), \sigma_1(s_1)) \in \overline{T}(B)$ by the monotonicity of $\overline{T}$. But then B is a bisimulation, again by Theorem 3.9.

Finally, for part 2, note that it is an immediate consequence of part 1 that $\leftrightarrows$, being the union of all bisimulations between $\mathbb{S}_0$ and $\mathbb{S}_1$, is a bisimulation itself. Hence, by definition, it is the greatest bisimulation between $\mathbb{S}_0$ and $\mathbb{S}_1$. In fact, it follows by the Knaster-Tarski theorem (on fixed points of monotone operations on complete lattices), that $\leftrightarrows$ is in fact the greatest fixed point of the map $\Lambda : R \mapsto \{(s_0, s_1) \mid (\sigma_0(s_0), \sigma_1(s_1)) \in \overline{T}(R)\}$. QED

In the case of Kripke polynomial functors, relation lifting can be characterized using *induction* on the construction of the functor.

Proposition 3.11 *Let S and S' be two sets, and let $R \subseteq S \times S'$ be a binary relation between S and S'. Then the following induction defines the relation lifting $\overline{K}(R) \subseteq KS \times KS'$, for each Kripke polynomial functor K:*

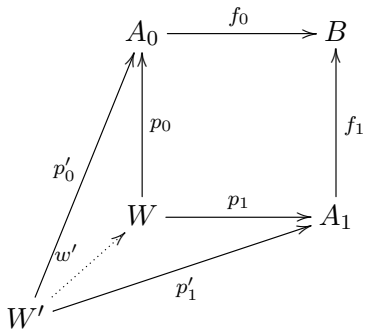
$$\begin{array}{rcl}
\overline{Id}(R) & := & R, \\
\overline{K_C}(R) & := & \Delta_C, \\
\overline{K_0 \times K_1}(R) & := & \{((x_0, x_1), (x'_0, x'_1)) \mid (x_0, x'_0) \in \overline{K_0}(R) \text{ and } (x_1, x'_1) \in \overline{K_1}(R)\}, \\
\overline{K_0 + K_1}(R) & := & \{(\kappa_0 x_0, \kappa_0 x'_0) \mid (x_0, x'_0) \in \overline{K_0}(R)\} \cup \{(\kappa_1 x_1, \kappa_1 x'_1) \mid (x_1, x'_1) \in \overline{K_1}(R)\}, \\
\overline{K^D}(R) & := & \{(f, f') \mid (f(d), f'(d)) \in \overline{K}(R) \text{ for all } d \in D\}, \\
\overline{PK}(R) & := & \{(Q, Q') \mid \forall q \in Q\, \exists q' \in Q'.\, (q, q') \in \overline{K}(R) \text{ and } \forall q' \in Q'\, \exists q \in Q.\, (q, q') \in \overline{K}(R)\}.
\end{array}$$

Here κ_0 and κ_1 are the co-projection maps associated with the coproduct, cf. Definition B.14.

3.3 Bisimilarity and behavioural equivalence: smooth functors

In Example 3.2 we saw that bisimilarity is a strictly stronger notion than behavioural equivalence. Here is a constraint on the functor that guarantees the two notions to coincide.

Definition 3.12 A *weak pullback* of two arrows $f_0 : A_0 \to B$, $f_1 : A_1 \to B$ in a category C is a pair of arrows $p_0 : W \to A_0$, $p_1 : W \to A_1$ such that (i) $f_0 \circ p_0 = f_1 \circ p_1$, while (ii) for every pair $p_0' : W' \to A_0$, $p_1' : W' \to A_1$ that also satisfies $f_0 \circ p_0' = f_1 \circ p_1'$, there is a mediating arrow $w' : W' \to W$ such that $p_0 \circ w' = p_0'$ and $p_1 \circ w' = p_1'$.
We will call a functor $T : \mathsf{C} \to \mathsf{C}'$ *smooth* if it preserves weak pullbacks; that is, if for any weak pullback (p_0, p_1) of any (f_0, f_1) in C, the pair (Tp_0, Tp_1) is a weak pullback of (Tf_0, Tf_1) in C'. $\lhd$



Note that the mediating arrow w' need not be unique: adding this requirement to the definition would give the more familiar, and stronger, notion of a *pullback*. The category Set has pullbacks: for $f_0 : A_0 \to B$ and $f_1 : A_1 \to B$, we can take the projections to A_0 and A_1 from the set $\mathsf{pb}(f_0, f_1) := \{(a_0, a_1) \in A_0 \times A_1 \mid f_0(a_0) = f_1(a_1)\}$.

Many but not all endofunctors on Set in fact preserve weak pullbacks.

Proposition 3.13 *All polynomial functors preserve pullbacks, and all Kripke polynomial functors preserve weak pullbacks. Furthermore, if T preserves weak pullbacks, then so does T_{Q}, for any set Q of proposition letters.*

The main reason that this prima facie rather exotic property is in fact of great importance in the theory of universal coalgebra, is the following fact (see also Fact B.36 in Appendix B).

Fact 3.14 *For any set functor T the following are equivalent:*
(1) T is smooth;
(2) $\overline{T}(R\,;Q) = \overline{T}R\,;\overline{T}Q$, for all pairs of relations $R \subseteq X \times Y$ and $Q \subseteq Y \times Z$.

Theorem 3.15 *If T is a smooth set functor, the following hold on the class of T-coalgebras:*
(1) the relational composition of two bisimulations is again a bisimulation;
(2) the notions of bisimilarity and behavioral equivalence coincide:

$$\simeq_T \;=\; \leftrightarroweq_T.$$

Proof. For the proof of the first statement, let $\mathbb{S}_0, \mathbb{S}_1$ and $\mathbb{S}_2$ be T-coalgebras, and let $B' : \mathbb{S}_0 \leftrightarroweq \mathbb{S}_1$ and $B'' : \mathbb{S}_1 \leftrightarroweq \mathbb{S}_2$ be bisimulations. We will show that $B := B'; B''$ is a bisimulation as well, between $\mathbb{S}_0$ and $\mathbb{S}_2$. For that purpose, take an arbitrary pair $(s_0, s_2) \in B$; then by definition there must be a state $s_1 \in S_1$ such that $(s_0, s_1) \in B'$ and $(s_1, s_2) \in B''$. But since B' and B'' are bisimulations, this means that $(\sigma_0 s_0, \sigma_1 s_1) \in \overline{T}B'$ and $(\sigma_1 s_1, \sigma_2 s_2) \in \overline{T}B''$, so that $(\sigma_0 s_0, \sigma_2 s_2) \in \overline{T}B'; \overline{T}B''$. Hence by smoothness we find that $(\sigma_0 s_0, \sigma_2 s_2) \in \overline{T}(B'; B'') = \overline{T}B$, as required.

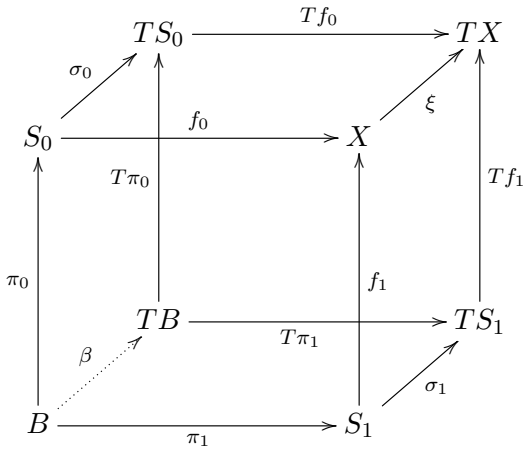
Turning to the second statement, let s_0 and s_1 be states of the T-coalgebras $\mathbb{S}_0$ and $\mathbb{S}_1$, respectively. We need to prove that $\mathbb{S}_0, s_0 \leftrightarrows \mathbb{S}_1, s_1$ iff $\mathbb{S}_0, s_0 \simeq \mathbb{S}_1, s_1$. Because of Proposition 3.1 it suffices to prove the direction from right to left.

Let $f_0 : \mathbb{S}_0 \to \mathbb{X}$ and $f_1 : \mathbb{S}_1 \to \mathbb{X}$ be two coalgebra morphisms such that $f_0(s_0) = f_1(s_1)$. Then in Set, the set $B := \{(s_0, s_1) \in S_0 \times S_1 \mid f_0(s_0) = f_1(s_1)\}$, together with the projection functions $\pi_0 : B \to S_0$ and $\pi_1 : B \to S_1$ constitutes a pullback of f_0 and f_1, cf. the square in the foreground of the picture. Because T preserves weak pullbacks, the square in the background of the picture is then a weak pullback diagram in Set.

Now consider the two arrows $\sigma_i \circ \pi_i : B \to T(S_i)$. First observe that $Tf_i \circ \sigma_i = \xi \circ f_i$ for each i, because each f_i is a coalgebra morphism. Hence, chasing the diagram we find that

$$\begin{aligned} Tf_0 \circ \sigma_0 \circ \pi_0 &= \xi \circ f_0 \circ \pi_0 \\ &= \xi \circ f_1 \circ \pi_1 = Tf_1 \circ \sigma_1 \circ \pi_1. \end{aligned}$$

Since $T\pi_0$ and $T\pi_1$ form a weak pullback of Tf_0 and Tf_1, this implies the existence of a mediating function $\beta : B \to TB$ such that $T\pi_i \circ \beta = \sigma_i \circ \pi_i$. In other words, $\mathbb{B} := (B, \beta)$ is a T-coalgebra, and the projection maps π_0 and π_1 are morphisms from $\mathbb{B}$ to $\mathbb{S}_0$ and $\mathbb{S}_1$, respectively.



QED

Corollary 3.16 *Let T be a smooth functor. Then for any T-coalgebra $\mathbb{S}$, the maximal bisimulation $\leftrightarrows$ on $\mathbb{S}$ is an equivalence relation.*

Proof. It is not hard to see that the maximal bisimulation $\leftrightarrows$ is always reflexive and symmetric. For transitivity, it follows by smoothness of T that the composition $\leftrightarrows ; \leftrightarrows$ is again a bisimulation; but since $\leftrightarrows$ is maximal we find that $\leftrightarrows ; \leftrightarrows \subseteq \leftrightarrows$, which is just another formulation of the transitivity of $\leftrightarrows$. QED

4 Covarieties

In universal *algebra* an important part is played by varieties: classes of algebras that are closed under the operations of taking homomorphic images, subalgebras and products of algebras. In this chapter, we introduce the notion of a *covariety* as a natural coalgebraic analog of a variety, and we consider some natural closure operations on classes of coalgebras.

4.1 Homomorphic images

Definition 4.1 Let T be some endofunctor on Set. If $f : \mathbb{S} \to \mathbb{S}'$ is a surjective coalgebra morphism between the T-coalgebras $\mathbb{S}$ and $\mathbb{S}'$, then we say that $\mathbb{S}'$ is a *homomorphic image* of $\mathbb{S}$, and write $\mathbb{S} \twoheadrightarrow \mathbb{S}'$, or $f : \mathbb{S} \twoheadrightarrow \mathbb{S}'$ if we want to make the morphism explicit. $\lhd$

In universal algebra, one finds a one-one correspondence between homomorphic images and congruences: special equivalence relations that interact well with the algebra structure. There is a similar connection here between homomorphic images and bisimulation equivalences: bisimulations that are also equivalence relations. The correspondence, however, is perfect only in the case of smooth functors.

Proposition 4.2 *Let $\mathbb{S} = (S, \sigma)$ be a T-coalgebra for some set functor T.*

(1) Given a bisimulation equivalence[6] *E on $\mathbb{S}$, there is a unique coalgebra structure $\overline{\sigma}$ on S/E such that the quotient map $q : S \to S/E$ is a coalgebra morphism.*

(2) If T preserves weak pullbacks, then the relation $\ker(f) := \{(s,t) \in S^2 \mid fs = ft\}$ is a bisimulation equivalence, for any coalgebra morphism $f : \mathbb{S} \to \mathbb{S}'$.

Proof. For part (1), we leave it as an exercise for the reader to show that the set $\overline{S} = S/E$ of E-cells, together with the quotient map q, is a coequalizer of the projection maps $\pi_0, \pi_1 : E \to S$:

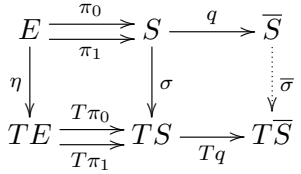
$$E \underset{\pi_1}{\overset{\pi_0}{\rightrightarrows}} S \xrightarrow{q} \overline{S}$$

Now assume that, next to being an equivalence relation, E is also a bisimulation on $\mathbb{S}$. Then by definition there is a coalgebra map $\eta : E \to TE$ such that both π_i are coalgebra morphisms $\pi_i : (E, \eta) \to \mathbb{S}$. It follows that

$$\begin{array}{rcll} Tq \circ \sigma \circ \pi_0 & = & Tq \circ T\pi_0 \circ \eta & (\pi_0 \text{ is a morphism}) \\ & = & T(q \circ \pi_0) \circ \eta & (\text{functoriality}) \\ & = & T(q \circ \pi_1) \circ \eta & (q \text{ is a coequalizer}) \\ & = & Tq \circ T\pi_1 \circ \eta & (\text{functoriality}) \\ & = & Tq \circ \sigma \circ \pi_1 & (\pi_1 \text{ is a morphism}) \end{array}$$

[6] A *bisimulation equivalence* is a bisimulation that is also an equivalence relation.

In other words, the map $Tq \circ \sigma : S \to T\overline{S}$ is a competitor for the coequalizer map q, and so there is a unique map $\overline{\sigma} : \overline{S} \to T\overline{S}$ such that $\overline{\sigma} \circ q = Tq \circ \sigma$, in a diagram:



Clearly then $\overline{\sigma}$ is the required coalgebra map on $\overline{S}$.

For the second part of the proposition, observe that $\ker(f)$ is the relational composition of the graph of f with its converse. But since the graph of a coalgebra morphism is a bisimulation (Statement (19) in Example 3.4), the result then follows from Theorem 3.15. QED

4.2 Subcoalgebras

The next class operation that we consider is that of taking subcoalgebras.

Definition 4.3 Let $\mathbb{X} = (X, \xi)$ be some T-coalgebra. We say that a subset S of X is *open* with respect to $\mathbb{X}$ if there is a coalgebra map $\sigma : S \to TS$ such that the inclusion map $\iota : S \to X$ is a coalgebra morphism from (S, σ) to (X, ξ); in this case we call the structure (S, σ) a *subcoalgebra* of $\mathbb{X}$, writing $\mathbb{S} \leq \mathbb{X}$. $\lhd$

Interestingly enough, the transition map of a subcoalgebra is completely determined by the underlying open set.

Proposition 4.4 *Let $\mathbb{S}_0 = (S, \sigma_0)$ and $\mathbb{S}_1 = (S, \sigma_1)$ be two subcoalgebras of the coalgebra $\mathbb{X}$. Then $\sigma_0 = \sigma_1$.*

Proof. The case of S being empty is trivial, so suppose otherwise. Then from the assumption that $\mathbb{S}_0$ and $\mathbb{S}_1$ are subcoalgebras of $\mathbb{A}$, we may infer that $(T\iota) \circ \sigma_0 = \xi \circ \iota = (T\iota) \circ \sigma_1$, where ι is the inclusion map of S into X. It follows from the functoriality of T that $T\iota$ is an injection, so that we may conclude that $\sigma_0 = \sigma_1$. QED

Remark 4.5 In case T is a standard functor, we may simplify Definition 4.3: in this case any subcoalgebra of $\mathbb{X} = (X, \xi)$ must be of the form $\mathbb{S} = (S, \xi \restriction_S)$ with $S \subseteq X$ and $\xi(s) \in TS$, for all $s \in S$.

Example 4.6 (a) Given a Kripke frame (S, R) we call a structure (S', R') a *subframe* of $\mathbb{S}$ if $S' \subseteq S$ and $R' = R \cap (S' \times S')$. Such a subframe is *generated* if its universe S' is closed under the relation R, that is, if $s \in S'$ and $(s, t) \in R$ imply $t \in S'$. It is easy to see that the generated subframes of a frame correspond to the subcoalgebras of its coalgebraic representation.

(b) Similarly, a state b in a deterministic finite automaton $\mathbb{A} = (A, \delta, F)$ is *reachable* from a state $a \in A$ if there is some word u such that $\widehat{\delta}(a, u) = B$. Given some automaton $\mathbb{A}$, the states that are reachable from a given state a form an open set.

(c) The empty set is always open: every coalgebra has the empty coalgebra as a subcoalgebra.

Some further observations concerning subcoalgebras are in order. First of all, the topological terminology is justified by the following proposition.

Proposition 4.7 *Given a coalgebra $\mathbb{X}$ for some set functor T, the collection $\tau_{\mathbb{X}}$ of $\mathbb{X}$-open sets forms a topology.*

Proof. Closure of $\tau_{\mathbb{X}}$ under taking (arbitrary) unions follows from Theorem **??**, together with the observation that

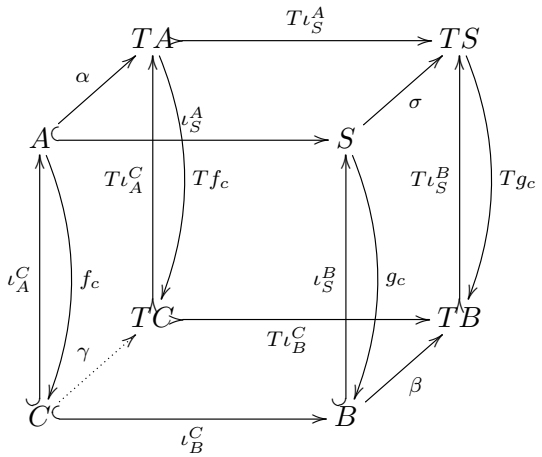
$$S \subseteq X \text{ is open with respect to } \mathbb{X} \text{ iff } \Delta_S \text{ is a bisimulation on } \mathbb{X}, \tag{24}$$

which in its turn is an immediate consequence of (19).

To prove that $\tau_{\mathbb{X}}$ is closed under taking finite intersections, assume that $\mathbb{A} = (A, \alpha)$ and $\mathbb{B} = (B, \beta)$ are two subcoalgebras of $\mathbb{S}$. The case where $A \cap B = \varnothing$ is trivial, so assume otherwise. Define $C := A \cap B$, fix some element $c \in C$, and consider the maps $f_c : A \to C$ and $g_c : S \to B$, given by

$$f_c(a) := \begin{cases} a & \text{if } a \in C \\ c & \text{otherwise} \end{cases} \qquad \text{and} \qquad g_c(s) := \begin{cases} s & \text{if } s \in B \\ c & \text{otherwise.} \end{cases}$$

Now consider the following diagram:



We define the map $\gamma : C \to TC$ by putting

$$\gamma := Tf_c \circ \alpha \circ \iota_A^C,$$

and claim that the resulting structure (C, γ) is a subcoalgebra of $\mathbb{S} = (S, \sigma)$. For this purpose we have to show that

$$\sigma \circ \iota_S^C = T\iota_S^C \circ \gamma. \tag{25}$$

We leave it as an exercise for the reader to show (25), using equalities such as $\iota_B^C \circ f_c = g_c \circ \iota_S^A$ and $g_c \circ \iota_S^B = \mathsf{id}_B$. QED

Remark 4.8 In the case where T is standard, the proof that $\tau_{\mathbb{X}}$ is closed under taking finite intersections can again be simplified. In order to prove that $C := A \cap B$ is the carrier of a subcoalgebra of $\mathbb{S}$, it suffices to show that $\sigma(c) \in TC$ for an arbitrary element c of $C := A \cap B$. But since T is standard we have $TC = TA \cap TB$, and we find $\sigma(c) \in TA \cap TB$ by the assumptions on A and B, respectively.

It follows from Proposition 4.7 that the partial order $(\tau_{\mathbb{X}}, \subseteq)$, consisting of the collection of open subsets of X ordered by set inclusion actually forms a *complete lattice*. In this lattice, the *join* of a collection $\mathcal{U}$ of opens is simply given as its union, but its *meet*, that is, the largest open set contained in each open $U \in \mathcal{U}$, is not necessarily its intersection.

In particular, given a subset S of X, there is an open set $V \subseteq X$ which is the *meet* of the collection $\{U \in \tau_{\mathbb{X}} \mid S \subseteq U\}$. However, there is no guarantee that V is the *intersection* of this collection, or, indeed, that S is actually a subset of V, as is witnessed by the example below. This example thus shows that it does not make sense to speak of the smallest subcoalgebra containing a given subset, or of *point-generated* subcoalgbras.

It follows from Proposition 4.7 that, given a subset S of (the carrier of) a coalgebra $\mathbb{X}$, there is a largest subcoalgebra $\langle S \rangle_{\mathbb{X}}$ of $\mathbb{X}$ (of which the carrier is) contained in S: Its carrier is given as the union of all open subsets of S. Note however, that this carrier might be empty, even if S is not.

Example 4.9 Consider the standard Euclidean topology on the real numbers, seen as a coalgebra for the *filter functor* F. This functor is a subfunctor of the (monotone) neighborhood functor which maps a set S to the collection of all *filters* on S[7] and a function $f : S \to S'$ simply to the function $Mf = \breve{P}\breve{P}f$. Prime examples of F-coalgebras are the topological spaces. To see this, represent a topology σ on the set S by the function mapping a point $s \in S$ to the collection $\{U \in \sigma \mid s \in U\}$ of its neighborhoods.

One can show that a set S of reals is open in the topological sense iff it is open in the sense of Definition 4.3 — in fact, this holds for any topology. Now take an arbitrary point r in $\mathbb{R}$. Obviously, the *meet* of all open neighborhoods containing r is the empty set. This example thus shows that in general it does not make sense to speak of point-generated subcoalgebras.

Before we turn to further coalgebraic constructions, consider the following natural link between homomorphic images and subcoalgebras.

Proposition 4.10 *Given a coalgebra morphism $f : \mathbb{S} \to \mathbb{S}'$, there is a (unique) subcoalgebra $f[\mathbb{S}]$ of $\mathbb{S}'$ such that $f : \mathbb{S} \to f[\mathbb{S}]$ is a surjective morphism.*

Proof. For a proof of this proposition, let $X := f[S]$ be the (set-theoretic) image of S under f, and let $g : X \to S$ be a right inverse of f, that is, $f(g(x)) = x$ for all $x \in X$. Now define $\xi : X \to TX$ by $\xi := Tf \circ \sigma \circ g$. It can be shown that the resulting structure $\mathbb{X}$ is always a subcoalgebra of $\mathbb{S}'$, and that $f : \mathbb{S} \to \mathbb{X}$ is a surjective morphism; further details are left for the reader. QED

[7] Recall that a filter on S is a collection $\mathcal{F}$ of subsets of S which is not only upward closed (with respect to the inclusion relation), but also closed under taking (finite) intersections, that is, $X \cap Y \in \mathcal{F}$ if $X, Y \in \mathcal{F}$.

4.3 Sums and other colimits

Our last example of a coalgebraic construction concerns the straightforward generalization of the disjoint union of Kripke models and frames. The idea is embodied in the following Proposition.

Proposition 4.11 *Let $\mathbb{S}_0 = (S_0, \sigma_0)$ and $\mathbb{S}_1 = (S_1, \sigma_1)$ be two T-coalgebras, and let $S := S_0 \uplus S_1$ be the disjoint union of S_0 and S_1. Then there is a unique arrow $\sigma : S \to TS$ making the embeddings κ_i into coalgebra morphisms.*

Proof. Consider the diagram below, where S, together with the embedding maps κ_0 and κ_1, is the coproduct of S_0 and S_1. Since the maps $T\kappa_0 \circ \sigma_0$ and $T\kappa_1 \circ \sigma_1$ provide an alternative co-cone, there must be a (unique) mediating arrow $\sigma : S \to TS$, making the two rectangles in the diagram commute.

$$\begin{array}{ccccc} S_0 & \xrightarrow{\kappa_0} & S & \xleftarrow{\kappa_1} & S_1 \\ {\scriptstyle\sigma_0}\downarrow & & \downarrow{\scriptstyle\sigma} & & \downarrow{\scriptstyle\sigma_1} \\ TS_0 & \xrightarrow[T\kappa_0]{} & TS & \xleftarrow[T\kappa_1]{} & TS_1 \end{array} \tag{26}$$

Clearly then this σ meets the requirements stated in the Proposition. QED

The sum of an arbitrary collection of coalgebras is defined as follows.

Definition 4.12 The *sum* $\coprod_I \mathbb{S}_i$ of a family $\{\mathbb{S}_i \mid i \in I\}$ of coalgebras for some set functor T, is defined by endowing the disjoint union $S := \biguplus_I S_i$ with the unique map $\sigma : S \to TS$ which turns all embeddings $\kappa_i : S_i \to S$ into coalgebra morphisms. $\lhd$

Sums are coproducts in the category of coalgebras.

Proposition 4.13 *Let $\{\mathbb{S}_i \mid i \in I\}$ be a family of coalgebras for some set functor T. Then $\coprod_I \mathbb{S}_i$ is a coproduct of $\{\mathbb{S}_i \mid i \in I\}$ in the category $\mathsf{Coalg}(T)$.*

In particular, Proposition 4.13 implies that coproducts actually *exist* (in the category of coalgebras for a given set functor T).

Remark 4.14 This result can be in fact be generalised to *arbitrary* colimits: it can be shown that for every set functor T, the category $\mathsf{Coalg}(T)$ has all colimits, and that these colimits are in fact based on the underlying colimits n Set.

4.4 Covarieties

We have now gathered all the basic class operations needed to define the notion of a covariety.

Definition 4.15 Let T be some endofunctor on Set. A class of T-coalgebras is a *covariety* if it is closed under taking homomorphic images, subcoalgebras and sums. The smallest covariety containing a class K of T-coalgebras is called the covariety *generated* by K, notation: $\mathsf{Covar}(\mathsf{K})$. $\lhd$

As in the case of universal algebra, in order to obtain a more succinct characterization of the covariety generated by a class of coalgebras, one may develop a calculus of class operations.

Definition 4.16 Let H, S and Σ denote the class operations of taking (isomorphic copies of) homomorphic images, subcoalgebras, and sums, respectively. $\lhd$

On the basis of these (and other) operations one may investigate the validity of 'inequalities' like $\mathsf{HS} \leq \mathsf{SH}$ (meaning that $\mathsf{HS}(\mathsf{K}) \subseteq \mathsf{SH}(\mathsf{K})$ for all classes K of coalgebras). We first prove that each of the three operations H, S and Σ is a *closure operation*; in each case the proof of this observation is straightforward.

Proposition 4.17 *Let* C *be one of the class operations* H, S *or* Σ, *and let* K *and* K' *be arbitrary classes of* T*-coalgebras. Then we find*

1) $\mathsf{K} \subseteq \mathsf{C}(\mathsf{K})$*;*

2) $\mathsf{K} \subseteq \mathsf{K}'$ *implies* $\mathsf{C}(\mathsf{K}) \subseteq \mathsf{C}(\mathsf{K}')$*;*

3) $\mathsf{C}(\mathsf{C}(\mathsf{K})) \subseteq \mathsf{C}(\mathsf{K})$.

In the following proposition we list some relevant inequalities involving class operations.

Proposition 4.18 *Let* K *be an arbitrary class of* T*-coalgebras. Then we find*

1) $\mathsf{HS}(\mathsf{K}) \subseteq \mathsf{SH}(\mathsf{K})$*;*

2) $\Sigma\mathsf{S}(\mathsf{K}) = \mathsf{S}\Sigma(\mathsf{K})$*;*

3) $\Sigma\mathsf{H}(\mathsf{K}) \subseteq \mathsf{H}\Sigma(\mathsf{K})$.

Proof. For part (1), let $\mathbb{A}$ be a coalgebra in $\mathsf{HS}(\mathsf{K})$. Then there is a coalgebra $\mathbb{C}$ in K, and a coalgebra $\mathbb{B}$, with morphisms e and f such that $e : \mathbb{B} \rightarrowtail \mathbb{C}$ and $f : \mathbb{B} \twoheadrightarrow \mathbb{A}$. Now consider the following push-out diagram in $\mathsf{Coalg}(T)$ (cf. Example B.20):

$$\begin{array}{ccc} \mathbb{B} & \overset{f}{\twoheadrightarrow} & \mathbb{A} \\ {\scriptstyle e}\downarrow & & \downarrow{\scriptstyle p_A} \\ \mathbb{C} & \underset{p_C}{\longrightarrow} & \mathbb{P} \end{array}$$

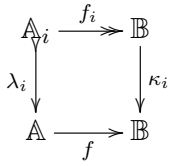
Since (P, p_A, p_C) is also the push-out of (B, e, f) in Set, it follows that the maps p_A and p_C are, respectively, surjective and injective.[8] But then we immediately find that $\mathbb{P} \in \mathsf{H}(\mathsf{K})$ and $\mathbb{A} \in \mathsf{SH}(\mathsf{K})$, as required.

We leave the proof of part (2) as an exercise for the reader. For part (3), assume that we have $\mathbb{B} \in \Sigma\mathsf{H}(\mathsf{K})$. That is, there is a family of algebras $\{\mathbb{A}_i \mid i \in I\}$ in K, and a family of surjective morphisms $\{f_i : \mathbb{A}_i \twoheadrightarrow \mathbb{B}_i \mid i \in I\}$ such that $\mathbb{B} = \Sigma_i \mathbb{A}_i$, with the injections $\kappa_i : \mathbb{B}_i \rightarrowtail \mathbb{B}$.

Let $\mathbb{A}$ be the sum of the coalgebras $\mathbb{A}_i$, with the morphisms $\lambda_i : \mathbb{A}_i \rightarrowtail \mathbb{A}$. Then the coalgebra $\mathbb{B}$, with the morphisms $\{\kappa_i \circ f_i : \mathbb{A}_i \to \mathbb{B}_i\}$, is a competitor cone of $\mathbb{A}$, and so there

[8] Some justification is needed here.

is a unique morphism $f : \mathbb{A} \to \mathbb{B}$ such that $f \circ \lambda_i = \kappa_i \circ f_i$, for all i:



It remains to show that f is surjective. For that purpose, take an arbitrary state $b \in \mathbb{B}$, then b must be of the form $b = \kappa_i(b_i)$ for some $b_i \in B_i$. By surjectivity of f_i there is a state $a \in A_i$ such that $b_i = f_i a_i$. But then we find that $b = \kappa_i f_i a_i = f \lambda_i a_i$, which obviously implies that b lies in the range of f. So f is surjective indeed. QED

The above results easily lead to the following coalgebraic analog of Tarski's HSP-theorem in universal algebra.

Theorem 4.19 *Let* K *be a class of* T*-coalgebras for some set functor* T*. Then*

$$\mathsf{Covar}(\mathsf{K}) = \mathsf{SH\Sigma}(\mathsf{K}).$$

Proof. It is an immediate consequence of the definitions that $\mathsf{SH\Sigma}(\mathsf{K}) \subseteq \mathsf{Covar}(\mathsf{K})$. For the opposite inclusion it suffices to show that the class $\mathsf{SH\Sigma}(\mathsf{K})$ is closed under the operations S, H and $\Sigma(\mathsf{K})$, but this is rather obvious by the Propositions 4.17 and 4.18. For instance, we easily verify that $\Sigma(\mathsf{SH\Sigma}(\mathsf{K})) = \mathsf{S\Sigma H\Sigma}(\mathsf{K}) \subseteq \mathsf{SH\Sigma\Sigma}(\mathsf{K}) = \mathsf{SH\Sigma}(\mathsf{K})$. QED