BIRKHOFFS VARIETY THEOREM FOR COALGEBRAS

H. PETER GUMM

Abstract. Assuming only boundedness of the type functor, we give a syntactical description of coequations, and we prove a coalgebraic version of Birkhoff's variety theorem.

1. Introduction

State based systems play an important role in computer science. Finite automata are used as language recognizers, Kripke-Structures and many variations of labeled transition systems model concurrent behaviour. An important aspect of many such models is nondeterminism. Nondeterministic models represent a lack of predictability of the next system state. This unpredictability may originate in communication events between systems far apart from each other and not coordinated with a common clock, or it may result from systems made up of subprocesses and threads which are being scheduled by the operating system, dependent on current system load and requests.

It is common to all such state based systems that the state is important in the implementation, but not in the specification of the system. Rather, the user expects a certain behaviour, without insisting on how this behaviour is implemented. As an example, consider object oriented programs. The state of such systems is "private" and not observable by the user, only certain attributes and methods are "public" and thereby observable. In order to judge whether a system satisfies a specification, it is only the observable behavior that counts. In particular, two systems are supposed to be equivalent, if they show the same observable behaviour.

It is only through a rather abstract view on universal algebra, that it can be seen that the above systems are actually dual to universal algebra, they are instances of what has been termed *coalgebras*. This insight is due to H. Reichel [Rei95] and has been the starting point for a rapidly developing theory of coalgebras. Computer scientists have realized that many notions and results scattered in automata theory, in the theory of object oriented programming, in process theory, and in modal logic find a common framework in the emerging theory of "universal coalgebra".

As a mathematical exercise, by dualizing universal algebra, coalgebras had already been studied in the sixties. A straightforward dualization of classical universal algebra, however, merely yields coalgebras consisting of a set A together with a family of maps $f_A : A \to n \cdot A$, where $n \cdot A$ is an abbreviation for $\biguplus_{i<n} A$, the n-fold disjoint union of A. It is not clear whether such structures might be of much use, certainly, they will not be able to describe automata, transition systems, or object oriented programs.

The key to revealing the coalgebraic nature of the relevant systems is a more general viewpoint on universal algebras. A universal algebra is just a set A with a map $f^A : F(A) \to A$, where $F : \mathcal{S}et \to \mathcal{S}et$ is a functor on the category of sets. This definition extends the classical notion, since given any type $\Delta = (n_i)_{i \in I}$, we can define

$$F(X) := \biguplus_{i \in I} X^{n_i},$$

so a universal algebra of type Δ is just a set A together with a single map $f^A : F(A) \to A$.

Dualizing this more general concept, we obtain a notion of coalgebra that encompasses all the mentioned applications. For instance, an automaton with state set S, alphabet M, transition map $\delta : S \times M \to S$, and set of terminal states $T \subseteq S$ can be encoded into a single map

$$\alpha : S \to S^M \times \{0, 1\},$$

hence can be seen as a coalgebra for the functor $F(X) := X^M \times \{0, 1\}$. Similarly, a nondeterministic transition system where for every state $s \in S$ and every input $m \in M$ a certain set $\delta(s, m) \subseteq S$ of successor states are possible, can be coded as a map

$$\alpha : S \to \mathcal{P}(S)^M,$$

that is, as a coalgebra for the functor $F(X) := \mathcal{P}(X)^M$, where $\mathcal{P}$ is the powerset functor.

As is often the case with rather general definitions, there are instances and applications that arise somehow unexpectedly. For example, by choosing the filter functor $\mathcal{F} : \mathcal{S}et \to \mathcal{S}et$ which associates with every set X the set of $\mathcal{F}(X)$ of all filters on X, we find that topological spaces are just $\mathcal{F}$-coalgebras.

It is surprising that a rich structure theory of universal coalgebra can be developed from the seemingly abstract definition. The standard introduction to the subject has been by J. Rutten [Rut00], however, it should be noted that many of his proofs depend on a further condition imposed on the type functor F, it is supposed to *preserve weak pullbacks.*

Rutten introduced the notion of *covariety* and *cofree coalgebra* and showed that under a further "boundedness"-condition on F, covarieties could be represented by subcoalgebras of a certain cofree structure $\mathcal{S}_X$.

Under the very same hypotheses on F, covarieties were characterized in [GS98] by means of closure under sums, substructures and homomorphic images. Furthermore, we analyzed the conditions under which a pair $(\mathcal{U}, \mathcal{A})$ of a coalgebra $\mathcal{A}$ with subcoalgebra $\mathcal{U} \leq \mathcal{A}$ gives rise to a covariety. Finally, in [Gum99b], we showed how these results could be seen as a coalgebraic version of Birkhoff's theorem, when elements of the cofree coalgebra $\mathcal{S}_X$ were interpreted as "behaviour patterns", or "coequations".

When preparing lecture notes for an introductory course on coalgebras ([Gum99a]) we realized that with an appropriate reformulation of the notion of "boundedness", the results mentioned above could be proven without assuming on "weak pullback preservation".

A shortcoming of the notion of "behaviour" or "coequation" had always been the lack of any syntactical description, preventing anything resembling an equational calculus. It was only recently, in joint work again with T. Schröder [GS00b], and with the help of J. Adámek, that we came across results of Trnková [Trn69], which led us to a detailed description of bounded functors (see [GS00b]). As a result, coequations can now be seen as congruence classes of labelled trees, each tree representing the possible behaviour emanating from a given state.

In this note, we shall give a full account of the coalgebraic version of Birkhoff's theorem, including a syntactical description of coequations.

2. Preliminaries

We collect all definitions, which are needed in the sequel. Some basic results will be quoted. For proofs we shall refer to the literature.

2.1. **Type functors.** A *type functor* $F : \mathcal{S}et \to \mathcal{S}et$ associates with every set X a set $F(X)$ and with every map $f : X \to Y$ a map $F(f) : F(X) \to F(Y)$, so that

(i) $F(id_X) = id_{F(X)}$, and
(ii) $F(g \circ f) = F(g) \circ F(f)$ whenever $f : X \to Y$ and $g : Y \to Z$.

We shall assume in the sequel that F is *nontrivial*, in that $F(X) \neq \emptyset$ whenever $X \neq \emptyset$. If $X \neq \emptyset$ and $f : X \to Y$ is injective, then f has a left inverse. It follows that $F(f)$ has a left inverse too, so it is also injective. Any surjective map $g : X \twoheadrightarrow Y$ has a right inverse by the axiom of choice, hence $F(g)$ is surjective, too.

2.2. **Some examples of functors.** We shall discuss some examples of functors that will be used in this article.

- The *powerset functor* $\mathcal{P}$ associates to a set X the set of all its subsets $\mathcal{P}(X) := \{U \mid U \subseteq X\}$ and to a map $f : X \to Y$ the map $\mathcal{P}(f) : \mathcal{P}(X) \to \mathcal{P}(Y)$ which is given by $\mathcal{P}(f)(U) := f[U] := \{f(u) \mid u \in U\}$.
- A variation of this definition yields the *finite powerset functor* $\mathcal{P}_\omega$ which assigns to a set X the set $\mathcal{P}_\omega(X)$ of all finite subsets of X. On maps $f : X \to Y$ this functor is defined as before.

- For a fixed set M, the M-th power functor associates to a set X the set $X^M := \{\tau : M \to X\}$ of all maps from M to X. A map $f : X \to Y$ is associated to a map $(f)^M : X^M \to Y^M$ by defining $(f)^M(\tau) := f \circ \tau$.
- For a fixed set C, the constant functor k_C associates to every set X the fixed set C, and to every map $f : X \to Y$ the identity map id_C.
- New functors can be obtained from old ones by composition, sums and products. As an example, the functor $C \times (-)^M$, which assigns to a set X the set $C \times X^M$ and to a map $f : X \to Y$ the obvious map $id_C \times f \circ (-)$, is obtained as the cartesian product of k_C with the M-th power functor.

2.3. **Natural Transformations.** A *natural transformation* η from a type functor G to a type functor F consists of a map $\eta_X : G(X) \to F(X)$ for every set X, so that for every map $f : X \to Y$ the following diagram commutes. η is called *surjective*, if η_X is surjective whenever $X \neq \emptyset$.

$$
\begin{array}{ccccc}
X & \quad & G(X) & \xrightarrow[\eta_X]{} & F(X) \\
\downarrow{\scriptstyle f} & & {\scriptstyle G(f)}\downarrow & & \downarrow{\scriptstyle F(f)} \\
Y & & G(X) & \xrightarrow{\eta_Y} & F(X)
\end{array}
$$

2.4. **Coalgebras.** A *coalgebra* of type F (also called F-coalgebra) is a pair $\mathcal{A} = (A, \alpha_A)$ consisting of a set A and a map $\alpha : A \to F(A)$. We call A the *carrier set* and α_A the *structure map*. A *homomorphism* between coalgebras $\mathcal{A} = (A, \alpha_A)$ and $\mathcal{B} = (B, \alpha_B)$ is a map $\varphi : A \to B$ for which the following diagram commutes:

$$
\begin{array}{ccc}
A & \xrightarrow{\varphi} & B \\
{\scriptstyle \alpha_A}\downarrow & & \downarrow{\scriptstyle \alpha_B} \\
F(A) & \xrightarrow{F(\varphi)} & F(B)
\end{array}
$$

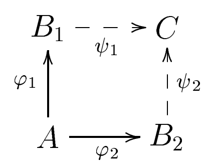
F-coalgebras with their homomorphisms form a category which will be denoted by $\mathcal{S}et_F$.

Epimorphisms in $\mathcal{S}et_F$ are just those homomorphisms which are surjective maps, but monomorphisms need not be injective, see [GS00a]. If φ is a bijective homomorphism, then its inverse φ^{-1} is a homomorphism too, so φ is an isomorphism (see [Rut00]). We write $\mathcal{A} \cong \mathcal{B}$ if they are *isomorphic*, i.e. if there exists an isomorphism between them.

2.5. **Colimits.** The forgetful functor from $\mathcal{S}et_F$ to $\mathcal{S}et$, associating with an F-Coalgebra $\mathcal{A} = (A, \alpha_A)$ its underlying set A, *creates colimits*, that is to say: Every colimit exists in $\mathcal{S}et_F$ and its underlying set is given by forming the colimit in $\mathcal{S}et$ and equipping it with the unique structure map, given by the colimit property (see [Rut00]). Two particular cases are of importance: *sums* and *pushouts*.

2.5.1. *Sums.* Given a family $(\mathcal{A}_i)_{i\in I}$ of coalgebras, their sum $\Sigma_{i\in I}\mathcal{A}_i$ is obtained by forming the disjoint union $\biguplus_{i\in I} A_i$ and equipping it with the unique structure map $\alpha : \biguplus_{i\in I} A_i \to F(\biguplus_{i\in I} A_i)$ which turns all embeddings $e_i : A_i \to \biguplus_{i\in I} A_i$ into homomorphisms.

2.5.2. *Pushouts.* Given coalgebras $\mathcal{A}$, $\mathcal{B}_1$, and $\mathcal{B}_2$ with homomorphisms $\varphi_1 : \mathcal{A} \to \mathcal{B}_1$ and $\varphi_2 : \mathcal{A} \to \mathcal{B}_2$, their *pushout* is obtained by first forming the pushout of φ_1 and φ_2 as set maps. This yields a set C with maps $\psi_1 : B_1 \to C$ and $\psi_2 : B_2 \to C$. On C there is a unique coalgebra structure $\alpha_C : C \to F(C)$ turning ψ_1 and ψ_2 into homomorphisms. It follows that $\mathcal{C} = (C, \alpha_C)$ is the pushout of the φ_i in $\mathcal{S}et_F$. If φ_1 is surjective, then so is ψ_2. If $A \neq \emptyset$ and φ_1 is injective, then it is a left-invertible in $\mathcal{S}et$. It follows that ψ_2 is left-invertible, in particular, injective.



We shall later need pushouts of an arbitrary collection $\varphi_i : \mathcal{A} \twoheadrightarrow \mathcal{B}_i$ of epimorphisms. By the same reasoning as above, this will yield us a coalgebra $\mathcal{C}$, together with a collection of epimorphisms $\psi_i : \mathcal{B}_i \twoheadrightarrow \mathcal{C}$.

2.6. **Subcoalgebras.** We write $\mathcal{U} \leq \mathcal{A}$, if $\mathcal{U} = (U, \alpha_U)$ is a *subcoalgebra* of $\mathcal{A}$, that is if $U \subseteq A$ and the natural inclusion map $\subseteq^A_U : U \to A$ is a homomorphism. Given a subset $U \subseteq A$, there is at most one structure map $\alpha_U : U \to F(U)$ making U the carrier set of a subcoalgebra of $\mathcal{A}$. Therefore, we use the term "subcoalgebra" both for $\mathcal{U}$ and for its carrier set U. Arbitrary unions and finite intersections of subcoalgebras are again subcoalgebras (see [GS00b]). If $\mathcal{U} \leq \mathcal{A}$ and $\varphi[U] \subseteq U$ for every homomorphism $\varphi : \mathcal{A} \to \mathcal{A}$, then we say that $\mathcal{U}$ is *invariant*.

2.7. **Homomorphic images.** If $\varphi : \mathcal{A} \twoheadrightarrow \mathcal{B}$ is a surjective homomorphism, then $\mathcal{B}$ is called a *homomorphic image* of $\mathcal{A}$. Any homomorphism $\varphi : \mathcal{A} \to \mathcal{B}$ factors as a $\mathcal{S}et$-map through its image $\varphi[A] = \{\varphi(a) \mid a \in A\}$. Let $\varphi = \subseteq^B_{\varphi[A]} \circ \varphi'$ be this factorization, and let $\psi : \varphi[A] \to A$ be a right-inverse to φ', then $\alpha' := F(\varphi') \circ \alpha_A \circ \psi$ uniquely defines a structure map on $\varphi[A]$ making it into a subcoalgebra $\varphi[\mathcal{A}] \leq \mathcal{B}$ and a homomorphic image of $\mathcal{A}$. Thus any homomorphism $\varphi : \mathcal{A} \to \mathcal{B}$ factors uniquely as $\mathcal{A} \twoheadrightarrow \varphi[\mathcal{A}] \leq \mathcal{B}$. Any surjective homomorphism $\varphi : \mathcal{A} \to \mathcal{B}$ induces a coalgebra structure, isomorphic to $\mathcal{B}$, on the quotient set $A/ker\varphi$. The *quotient map* $\pi_{ker\varphi}$ with $\pi_{ker\varphi}(a) = [a]ker\varphi$ is a surjective homomorphism.

2.8. **Conjunct sums.** A coalgebra $\mathcal{A}$ is a *conjunct sum* of a family $(\mathcal{A}_i)_{i\in I}$, provided that there is a family of embeddings (injective homomorphisms) $\varphi_i : \mathcal{A}_i \to \mathcal{A}$ so that $\mathcal{A}$ is the union of the subcoalgebras $\varphi_i[\mathcal{A}_i]$. Equivalently, there is a surjective homomorphism $\varphi : \Sigma_{i\in I}\mathcal{A}_i \to \mathcal{A}$, so that all compositions $\varphi \circ e_i : \mathcal{A}_i \to \mathcal{A}$ are injective.

Definition 2.1. *A coalgebra $\mathcal{A}$ is called* residually κ*, if it is a conjunct sum of coalgebras of cardinality at most κ. If $\mathcal{K}$ is a class of coalgebras, we shall denote with $\mathcal{K}_{\leq\kappa}$ the class of all members of $\mathcal{K}$ which are residually κ.*

2.9. **Bisimulations.** Bisimulations are the compatible relations between coalgebras. A relation $R \subseteq A_1 \times A_2$ between coalgebras $\mathcal{A}_1$ and $\mathcal{A}_2$ is a bisimulation, if R can be equipped with a coalgebra structure so that the projection maps $\pi_i : R \to A_i$ become homomorphisms.

The union of bisimulations is again a bisimulation, hence there is always a largest bisimulation, denoted by $\sim_{\mathcal{A}_1,\mathcal{A}_2}$ between any two coalgebras $\mathcal{A}_1$ and $\mathcal{A}_2$. We write $\sim_{\mathcal{A}}$ instead of $\sim_{\mathcal{A},\mathcal{A}}$.

2.10. **Examples.** We discuss three examples which are relevant in computer science. *Kripke models* are used to model nondeterministic behaviour of processes, and they provide a standard semantics for modal logic. *Deterministic automata* model "black boxes" whose internal states are hidden. Their behaviour can only be inferred from the output generated. When the output set is restricted to the 2-element set $\{True, False\}$, deterministic automata are used to define languages. *Nondeterministic automata* generalize both of the above structures.

2.10.1. *Kripke Models.* For a fixed set C of "propositions", a *Kripke model* is usually introduced as a set A of states together with a transition relation $R \subseteq A \times A$ and a valuation map $V : C \to \mathcal{P}(A)$. The latter assoiates to a proposition c the set of states in which c is to hold. Obviously, we can combine these data into a single map:

$$\alpha : A \to \mathcal{P}(C) \times \mathcal{P}(A),$$

hence a Kripke Model is just a coalgebra of type $k_{\mathcal{P}(C)} \times \mathcal{P}$, where $k_{\mathcal{P}(C)}$ is the constant functor with value $\mathcal{P}(C)$ and $\mathcal{P}$ the powerset functor.

2.10.2. *Deterministic Automata.* An *automaton* with *input set M* and *output set C* consists of a set A of *states*, and two maps

$$\begin{aligned} \gamma_A &: A \to C \\ \delta_A &: A \times M \to A. \end{aligned}$$

γ_A is called the *output map* and δ_A the *state transition* map. We can combine γ and δ into a single map

$$\alpha_A : A \to C \times A^M$$

by setting

$$\alpha_A(a) := (\gamma_A(a), \tau),$$

where

$$\forall m \in M.\ \tau(m) := \delta_A(a, m).$$

Thus automata with input set M and output set C are nothing but coalgebras for the functor $C \times (-)^M$.

It is straightforward to check that a coalgebra homomorphism φ between two $C \times (-)^M$-coalgebras $\mathcal{A}$ and $\mathcal{B}$ is the same as an automaton homomorphism, i.e. it is required to satisfy

$$\begin{aligned} \gamma_A(a) &= \gamma_B(\varphi(a)) \\ \varphi(\delta_A(a, m)) &= \delta_B(\varphi(a), m). \end{aligned}$$

In the sequel we shall drop the indices to α, δ and γ, whenever there is no danger of confusion.

Considering M as a (possibly infinite) *alphabet*, we let M^* denote the set of all words with letters from M. Let ε denote the empty word and for $m \in M$ and $w \in M^*$ let $m \cdot w$ be the word with first letter m and remaining word w.

Now δ can be extended to a map $\delta^* : A \times M^* \to A$ by inductively defining:

$$\begin{aligned} \delta^*(a, \varepsilon) &:= a, \text{and} \\ \delta^*(a, m \cdot w) &:= \delta^*(\delta(a, m), w). \end{aligned}$$

Then it is easy to see that a subset $U \subseteq A$ is a subcoalgebra of $\mathcal{A}$ iff it satisfies: $\forall u \in U. \forall w \in M^*. \delta^*(u, w) \in U$.

A bisimulation between two automata $\mathcal{A}$ and $\mathcal{B}$ is easily seen to be a relation $R \subseteq A \times B$ satisfying for all $(a, b) \in R$:

$$\gamma(a) = \gamma(b), \quad \text{and}$$

$$\forall m \in M.\ \delta(a, m)\ R\ \delta(b, m).$$

The largest bisimulation is given as

$$\sim_{\mathcal{A},\mathcal{B}} = \ \{(a, b) \in A \times B \mid \forall w \in M^*.\ \gamma(\delta^*(a, w)) = \gamma(\delta^*(b, w))\ \}.$$

In case $\mathcal{A} = \mathcal{B}$ and $C = \{0, 1\}$, this is also known as the *Nerode-congruence* of language theory. It is a coincidence that in this case $\sim_{\mathcal{A}}$ is an equivalence relation. In general coalgebras, $\sim_{\mathcal{A}}$ is reflexive and symmetric, but not necessarily transitive (see [GS00a]).

2.10.3. *Nondeterministic automata.* Nondeterministic automata with input set M and output set C are obtained as coalgebras for the functor that associates to a set X the set $C \times \mathcal{P}(X)^M$. Obviously, nondeterministic automata generalize both Kripke models and deterministic automata.

3. Covarieties, cofree coalgebras, and Birkhoff's theorem

It is our aim to specify classes of coalgebras in a way analogous to the equational specification of varieties of universal algebras given by Birkhoff's theorem.

3.1. **Covarieties.**

Covarieties will be introduced as classes closed under structural closure operators.

Definition 3.1. *Given a class $\mathcal{K}$ of F-coalgebras, we denote by*

- $\mathcal{I}(\mathcal{K})$ *the class of all isomorphic copies, by*
- $\mathcal{H}(\mathcal{K})$ *the class of all homomorphic images, by*
- $\mathcal{S}(\mathcal{K})$ *the class of all subcoalgebras, by*
- $\Sigma(\mathcal{K})$ *the class of all sums, and by*
- $\Sigma_C(\mathcal{K})$ *the class of all conjunct sums*

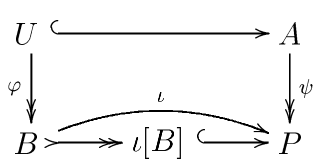
of coalgebras in $\mathcal{K}$. A class $\mathcal{K}$ of coalgebras is called a covariety*, if it is closed under the operators Σ, $\mathcal{H}$, and $\mathcal{S}$.*

Obviously, all operators are idempotent, up to isomorphism, in that $\mathcal{O}(\mathcal{O}(\mathcal{K})) \subseteq \mathcal{I}(\mathcal{O}(\mathcal{K}))$ for each of the operators introduced above and for all classes $\mathcal{K}$. The following lemma was shown in [GS98]. Even though, the article assumed that F should "preserve weak pullbacks", this assumption was not invoked in the original proof.

Lemma 3.2. *For an arbitrary class $\mathcal{K}$, we have*

(i) $\mathcal{I}(\mathcal{S}(\mathcal{K})) \subseteq \mathcal{S}(\mathcal{I}(\mathcal{K}))$,
(ii) $\mathcal{H}(\mathcal{S}(\mathcal{K})) \subseteq \mathcal{I}(\mathcal{S}(\mathcal{H}(\mathcal{K})))$,
(iii) $\Sigma(\mathcal{S}(\mathcal{K})) \subseteq \mathcal{S}(\Sigma(\mathcal{K}))$,
(iv) $\Sigma(\mathcal{H}(\mathcal{K})) \subseteq \mathcal{H}(\Sigma(\mathcal{K}))$.

There are two essential ingredients for the proof of this lemma. One is that every homomorphism $\varphi : \mathcal{A} \to \mathcal{B}$ factors through its image $\varphi[\mathcal{A}]$ which is a subcoalgebra of $\mathcal{B}$, and the other is the fact that the forgetful functor from $\mathcal{S}et_F$ to $\mathcal{S}et$ creates colimits, as dicussed in section 2.5. For instance, to prove (ii), we start with $\mathcal{U} \leq \mathcal{A}$ and $\varphi : \mathcal{U} \twoheadrightarrow \mathcal{B}$ where $\mathcal{A} \in \mathcal{K}$. Let $\mathcal{P}$ be the pushout of φ with the inclusion homomorphism $\subseteq^A_U : \mathcal{U} \to \mathcal{A}$. Then the remarks in section 2.5.2 imply that there is a surjective homomorphism $\psi : \mathcal{A} \twoheadrightarrow P$ and an injective homomorphism $\iota : \mathcal{B} \to \mathcal{P}$. Hence $\mathcal{B}$ is isomorphic to the subcoalgebra $\iota[\mathcal{B}]$ of $\mathcal{P}$.



From the lemma it follows immediately:

Theorem 3.3. $\mathcal{SH}\Sigma(\mathcal{K})$ *is the smallest covariety containing $\mathcal{K}$.*

Even though the above may seem to be perfectly dual to the universal algebraic situation, by switching the role of the operators $\mathcal{H}$ and $\mathcal{S}$ and by replacing products with sums, we caution the reader, that the situation is not that clear cut. For one thing, we are talking here about type functors F that are more general than the type functors $\Sigma_{i\in I}(-)^{n_i}$ of universal algebra. On the other hand, it was shown in [GS00b] that the intersection of finitely many subcoalgebras of a given coalgebra is again a subcoalgebra. From this it can be concluded that

$$\mathcal{S}\Sigma(\mathcal{K}) = \Sigma\mathcal{S}(\mathcal{K})$$

holds for arbitrary classes of coalgebras. The dual, in the above sense, of this result is false for universal algebras. We shall later need another consequence of the closure of subcoalgebras under finite intersections:

Lemma 3.4. *For an arbitrary class $\mathcal{K}$ of coalgebras*

$$\mathcal{S}\Sigma_C(\mathcal{K}) \subseteq \Sigma_C\mathcal{S}(\mathcal{K}).$$

Proof. If $\mathcal{A} \in \mathcal{S}\Sigma_C(\mathcal{K})$, then there is a coalgebra $\mathcal{C}$ and there are subcoalgebras $\mathcal{B}_i \leq \mathcal{C}$ so that $\mathcal{A} \leq \bigcup_{i\in I}\mathcal{B}_i \leq \mathcal{C}$, and each $\mathcal{B}_i \in \mathcal{K}$. By [GS00b], each $A \cap B_i$ is a subcoalgebra of $\mathcal{C}$, so $A = \bigcup_{i\in I}(A \cap B_i)$ is a conjunct sum of subcoalgebras of the $\mathcal{B}_i$. □

3.2. **Terminal Automata.** We consider an automaton with input set M and output set C as a black box. The states are not visible, but only the output that is generated via γ. We can enter a sequence $w \in M^*$ of inputs and observe the resulting output. In this way, every state s of such an automaton gives rise to a map $\omega(s) : M^* \to C$ defined by

$$\omega(s)(w) := \gamma(\delta^*(s, w)).$$

In an intuitive sense, the function $\omega(s)$ encodes every possible "observable behaviour" arising from state s. Hence we can view $T := C^{M^*}$ as the set of all possible behaviours of automata. T can itself be made into an automaton $\mathcal{T}$ by defining for any $\tau \in C^{M^*}$:

$$\begin{aligned} \gamma_T(\tau) &:= \tau(\varepsilon) \\ \delta_T(\tau, m) &:= \tau_m, \text{ where } \tau_m(w) = \tau(m \cdot w). \end{aligned}$$

It is now easy to verify that for every automaton $\mathcal{A}$ with input set M and output set C, the map ω is the unique homomorphism $\omega : \mathcal{A} \to \mathcal{T}$. Thus, $\mathcal{T}$ is *terminal* in the sense of the following section.

3.3. **Terminal Coalgebras.** An F-coalgebra $\mathcal{T} = (T, \alpha_T)$ is called *terminal* (or *final*), if for every $\mathcal{A} \in \mathcal{S}et_F$ there is a unique homomorphism $\omega : \mathcal{A} \to \mathcal{T}$. If the uniqueness requirement is dropped, then $\mathcal{T}$ is called *weakly terminal*.

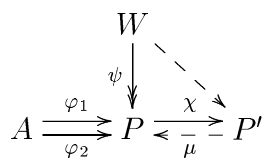
As in the above automata example, terminal coalgebras, if they exist, play an important role, since they embody all possible "behaviours" from a whole class of coalgebras. To be precise, for every coalgebra $\mathcal{A} \in \mathcal{S}et_F$ and any $a \in A$, there is precisely one element $\omega(a)$ of the terminal coalgebra which is bisimilar to a. In particular, no two elements of the terminal coalgebra are bisimilar, that is $\sim_{\mathcal{T}} = \Delta_T := \{(x,x) \mid x \in T\}$.

Given just a weakly terminal coalgebra, the terminal one can be constructed as the smallest homomorphic image:

Lemma 3.5. *Let $\mathcal{W}$ be weakly terminal and let $\mathcal{P}$ be the pushout of all quotients of $\mathcal{W}$. Then $\mathcal{P}$ is terminal.*

Proof. The pushout yields an epimorphism $\psi : W \twoheadrightarrow P$. Note that $\mathcal{P}$ is also weakly terminal. To see that it is indeed terminal, consider a parallel pair $\varphi_1, \varphi_2 : \mathcal{A} \to \mathcal{P}$ of homomorphisms. We must show that $\varphi_1 = \varphi_2$.

Consider their coequalizer $\chi : \mathcal{P} \to \mathcal{P}'$. It is surjective, for such is the case in $\mathcal{S}et$. Hence $\chi \circ \psi : \mathcal{W} \to \mathcal{P}'$ is an epimorphism. By construction of $\mathcal{P}$, there must be a homomorphism $\mu : \mathcal{P}' \to P$ with $\mu \circ \chi \circ \psi = \psi$. It follows that $\mu \circ \chi = id_P$, so $\varphi_1 = \mu \circ \chi \circ \varphi_1 = \mu \circ \chi \circ \varphi_2 = \varphi_2$.

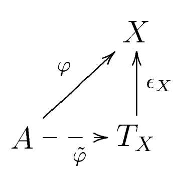


□

3.4. **Cofree coalgebras.** Our standard examples of coalgebras - automata and Kripke models - came equipped with a notion of "output". The possible outputs arising from a given state can be combined in a structure representing the "behaviour" arising in this state. In specifying classes of coalgebras, it is natural to define them by means of their permitted, or their forbidden behaviours.

For a general type functor $F : \mathcal{S}et \to \mathcal{S}et$, there may be no standard notion of "output" available, therefore, we consider what would happen, if we would adjoin a fixed set X of outputs:

Definition 3.6. *Let X be a set. An F-coalgebra $\mathcal{T}_X$ together with a $\mathcal{S}et$-map $\epsilon_X : T_X \to X$ is called* cofree over X, *if for every coalgebra $\mathcal{A}$ and every map $\varphi : A \to X$, there is exactly one homomorphism $\tilde{\varphi} : \mathcal{A} \to \mathcal{T}_X$ with $\epsilon_X \circ \tilde{\varphi} = \varphi$.*



An F-coalgebra $\mathcal{A} = (A, \alpha_A)$ together with a coloring $\varphi : A \to X$ is just a $k_X \times F$-coalgebra $\mathcal{A}' = (A, (\varphi, \alpha_A))$, and conversely. In the same way, the coalgebra $\mathcal{T}_X = (T_X, \alpha_{T_X})$ with coloring ϵ_X is cofree over X iff $\mathcal{T}_X' = (T_X, (\epsilon_X, \alpha_{T_X}))$ ist terminal as $k_X \times F$-coalgebra.

In this way, X may be seen as an added output set and an X-coloring as an output map. $\mathcal{T}_X$ is then the set of all behaviours with respect to this output set. Therefore, we shall call each element of T_X a "behaviour pattern".

Let $\mathcal{K}$ be a class of F-coalgebras, and let $\mathcal{T}_X$ be cofree over X. Every map $\varphi : \mathcal{A} \to X$ where $\mathcal{A} \in \mathcal{K}$, determines a subcoalgebra $\tilde{\varphi}[\mathcal{A}] \leq \mathcal{T}_X$. Put

$$\mathcal{T}_X(\mathcal{K}) := \bigcup \{\tilde{\varphi}[\mathcal{A}] \mid \mathcal{A} \in \mathcal{K}, \varphi : \mathcal{A} \to X\},$$

then this defines a subcoalgebra of $\mathcal{T}_X$ which is moreover *invariant*, in that it is closed under every endomorphism of $\mathcal{T}_X$. If $\mathcal{K}$ is closed under sums and homomorphic images, we even have $\mathcal{T}_X(\mathcal{K}) \in \mathcal{K}$.

Conversely, if U is a subcoalgebra of $\mathcal{T}_X$, then from [GS98] it follows that

$$\mathcal{Q}(U) := \{\mathcal{A} \in \mathcal{S}et_F \mid \forall \varphi : \mathcal{A} \to X. \tilde{\varphi}[A] \subseteq U\}$$

is a covariety. Only for showing closure under $\mathcal{S}$ will the cofreeness of $\mathcal{T}_X$ be needed.

It is now natural to combine the constructions. Let $\mathcal{V}_X(\mathcal{K}) := \mathcal{Q}(\mathcal{T}_X(\mathcal{K}))$, then

Lemma 3.7. *For any $\kappa \leq |X|$, if $\mathcal{T}_X$ exists, then*

$$\mathcal{V}_X(\mathcal{K})_{\leq \kappa} = \mathcal{SH}\Sigma(\mathcal{K})_{\leq \kappa}.$$

Proof. Since $\mathcal{V}_X(\mathcal{K})$ is a covariety containing $\mathcal{K}$, it must contain $\mathcal{SH}\Sigma(\mathcal{K})$. Conversely, let $\mathcal{A} \in \mathcal{V}_X(\mathcal{K})_{\leq \kappa}$. Then $\mathcal{A}$ is a conjunct sum of $\mathcal{A}_i$, each of size at most κ. It suffices to show that $\mathcal{A}_i \in \mathcal{S}\Sigma_C(\mathcal{K})$. Choose any injective map $\varphi : A_i \to X$, it extends to an injective homomorphism $\tilde{\varphi} : \mathcal{A}_i \to \mathcal{T}_X$. Consequently, $\mathcal{A}_i$ is isomorphic to a subcoalgebra of $\mathcal{T}_X(\mathcal{K}) \in \Sigma_C\mathcal{H}(\mathcal{K})$. Hence $\mathcal{A} \in \Sigma_C\mathcal{IS}\Sigma_C\mathcal{H}(\mathcal{K}) \subseteq \mathcal{H}\Sigma\mathcal{SH}\Sigma\mathcal{H}(\mathcal{K}) \subseteq \mathcal{SH}\Sigma(\mathcal{K})$. □

As a corollary to the proof and by invoking lemma 3.4 we get:

Corollary 3.8. *If the cofree F-coalgebras $\mathcal{T}_X$ exists for some $|X| \geq \kappa$, then $\mathcal{V}(\mathcal{K})_{\leq \kappa} = \Sigma_C\mathcal{SH}(\mathcal{K})_{\leq \kappa}$.*

4. Bounded Functors

It follows from an observation of Lambek (for a proof see [Rut00]), that the structure map $\alpha_T : T \to F(T)$ on a terminal coalgebra must be bijective. Consequently, a terminal coalgebra cannot exist, for instance, when F is the powerset functor. Fortunately, however, most functors of relevance in computer science do not grow in such an uncontrollable way, that is, we can put a bound on their growth:

Definition 4.1. *A functor $F : \mathcal{S}et \to \mathcal{S}et$ is* bounded, *if there is a set X such that for every F-coalgebra $\mathcal{A}$ and every $a \in A$ there is a subcoalgebra $\mathcal{U} \leq \mathcal{A}$ with $a \in U$ and $|U| \leq |X|$.*

Many functors F have the property that F-subcoalgebras are closed under arbitrary intersections. In this case, for every $\mathcal{A} \in \mathcal{S}et_F$ and every $a \in A$, there exists a smallest subcoalgebra $\langle a \rangle$ containing a. Such subcoalgebras are called "one-generated". Functors F, as described above, are therefore bounded, iff there is a cardinal bound $|X|$ on the size of one-generated F-subcoalgebras.

From this remark it follows that all of the previously mentioned functors, with the exception of $\mathcal{P}$, are bounded: Every one-generated subautomaton is of size at most $|M^*|$, and similarly, every $\mathcal{P}_\omega$-coalgebra has a substructure of size at most ω.

Obviously, F is bounded if and only if $\mathcal{S}et_F$ is residually small. Thus there is a *set* of generators, consisting of coalgebras of size at most κ, so with the help of the "special adjoint functor theorem", it can be shown that cofree F-coalgebras exist. For a straightforward proof, see [GS99].

At the same time, if F is bounded by X, we have $\mathcal{K}_{\leq |X|} = \mathcal{K}$ for any class $\mathcal{K}$ of F-coalgebras. Combining these, we have:

Theorem 4.2. *Let F be bounded by X, then the cofree coalgebra $\mathcal{T}_X$ exists and*

- *for every covariety $\mathcal{K}$, there exists a invariant subcoalgebra $\mathcal{U}$ of $\mathcal{T}_X$ with $\mathcal{Q}(U) = \mathcal{K}$,*
- *for every invariant subcoalgebra U of $\mathcal{T}_X$, there exists a covariety $\mathcal{K}$ with $U = \mathcal{T}_X(\mathcal{K})$.*

5. Coequations

Covarieties correspond uniquely to fully invariant subcoalgebras of a cofree coalgebra, hence it is natural to define a *coequation with covariables in* X as an element of the cofree coalgebra $\mathcal{T}_X$, i.e. as a behaviour pattern with variables in X.

Starting with a set $E \subseteq T_X$ of coequations, then a natural consequence relation is given by forming the largest fully invariant subcoalgebra contained in E. Then, however, an element $e \in E$ need not be in the set of consequences of E. In particular, it is unlikely that a single equation e will have any consequence.

For this reason, we have opted to define validity as *exclusion.* An element of $\mathcal{T}_X$ is a "forbidden behaviour pattern" and an element a of a coalgebra $\mathcal{A}$ satisfies/obeys the coequation, if the corresponding pattern is *avoided.* Such specifications by avoidance are common in many fields of mathematics. Planar graphs, for instance are defined by avoidance of two specific graphs, $K_{3,3}$ and K_5, modular, resp. distributive, lattices by the avoidance of the special lattices N_5, resp. N_5 and M_3. Therefore, we define for a coequation $e \in T_X$, any coalgebra $\mathcal{A}$, element $a \in A$ and any coloring $\varphi : A \to X$:

$$\begin{aligned} \mathcal{A}, a \models_\varphi e &:\iff \tilde{\varphi}(a) \neq e \\ \mathcal{A}, a \models e &:\iff \forall \varphi : A \to X.\ \mathcal{A}, a \models_\varphi e \\ \mathcal{A} \models e &:\iff \forall a \in A.\ \mathcal{A}, a \models e. \end{aligned}$$

For a set E of coequations, we define their model class, and for a class $\mathcal{K}$ of coalgebras their (avoided) behaviour patterns as

$$\begin{aligned} \mathcal{M}od(E) &:= \{\mathcal{A} \in \mathcal{S}et_F \mid \forall e \in E.\ \mathcal{A} \models e\}, \text{and} \\ \mathcal{B}eh_X(\mathcal{K}) &:= \{e \in T_X \mid \forall \mathcal{A} \in \mathcal{K}.\ \mathcal{K} \models e\}. \end{aligned}$$

Combining these definitions with lemma 3.7, we now obtain the following coalgebraic version of Birkhoff's theorem:

Theorem 5.1. *Let the type functor F be bounded by X. Then for any class $\mathcal{K}$ of F-coalgebras we have*

$$\mathcal{SH}\Sigma(\mathcal{K}) = \mathcal{M}od(\mathcal{B}eh_X(\mathcal{K})).$$

5.1. **Syntax.** In spite of the formal dual analogy to Birkhoff's theorem, one essential ingredient seems to be missing, that is the syntactical nature of coequations, and with it a syntactical style of reasoning. In universal algebra, the syntactical nature of equations is a direct consequence of the restricted nature of the type functor as $F = \Sigma_{i\in I}(-)^{n_i}$. So at first, the generality in the type functors permitted for coalgebras, seems to forbid any syntactical description. It was only after J. Adámek pointed us to some relevant work of Trnková, which in [GS00b], we were able to use for characterizing bounded functors as follows:

Theorem 5.2 ([GS00b])**.** *A functor F is bounded, iff there exist sets C and M and a surjective natural transformation $\eta : C \times (-)^M \to F$.*

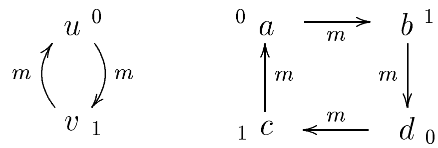
In the basic case, when $F = C \times (-)^M$, we have already seen that the cofree F-coalgebra over X has $T_X = (C \times X)^{M^*}$ as underlying set. Coequations are infinite trees with nodes labelled by pairs $(c, x) \in C \times X$, each node having precisely $|M|$ many sons.

In the general case, when F is bounded, there are sets C and M and a surjective natural transformation $\eta : C \times (-)^M \to F$. It follows that every F-coalgebra is of the form $\mathcal{A}_\eta = (A, \eta_A \circ \alpha_A)$ for some $C \times (-)^M$-coalgebra $\mathcal{A} = (A, \alpha_A)$. If $\mathcal{W}$ is (weakly) terminal as $C \times (-)^M$-coalgebra, then $\mathcal{W}_\eta$ is weakly terminal as F-coalgebra.

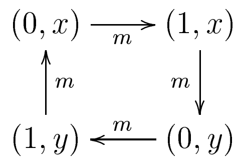
Consequently, the terminal F-coalgebra is a factor of $\mathcal{T}_\eta$ where $\mathcal{T}$ is the terminal $C \times (-)^M$-coalgebra. The latter is the set of all infinite M-branching trees with nodes labelled from C. Hence the terminal F-coalgebra consists of equivalence classes of infinite trees.

Quite similarly, the elements of the F-coalgebra cofree over X consists of equivalence classes of M-branching infinite trees with labels from $C \times X$. Hence every coequation is an equivalence class of some "syntactic object".

We shall give an example that shows the role of covariables in coequations. As mentioned before, they can be interpreted as an additional output by which previously indistinguishable states can be told apart. Let, for instance, $M = \{m\}$ and $C = \{0, 1\}$. If we want to distinguish the following two coalgebras,



where the indices at the states denote their respective γ-values, we shall need at least two covariables. By differently labelling nodes with the same output, we can enforce a distinction. The 2-variable coequation, for instance, which is obtained by unfolding the following diagram at node $(0, x)$, distinguishes the above coalgebras.



It is obviously avoided by the first one, but not by the second one, as witnessed by the coloring φ with $\varphi(a) = \varphi(b) = x$ and $\varphi(c) = \varphi(d) = y$.

5.2. **Coequational Implication.** For a coequation f and a set of coequations E, we define

$$E \models f$$

to mean that every coalgebra A satisfying all coequations in E must also satisfy f. Let us restrict ourself to the case of the functor $F = C \times (-)^M$. In this case, coequations are infinite $C \times X$-labelled trees, rather than equivalence classes of such trees. For any tree τ, a word $w \in M^*$ determines a subtree $\delta^*(\tau, w)$ of τ.

Theorem 5.3. *$E \models f$ if and only if there is an endomorphism $\varphi : \mathcal{T}_X \to T_X$ mapping a subtree of f into E.*

Proof. Suppose $E \models f$, then $\langle f \rangle := \{\delta^*(f, w) \mid w \in M^*\}$ is a subcoalgebra of $\mathcal{T}_X$, containing f. Thus $\langle f \rangle \not\models f$, hence $\langle f \rangle \not\models e$ for some $e \in E$, hence there is an element $a \in \langle f \rangle$ and a map $\varphi : \langle f \rangle \to X$ with $\tilde{\varphi}(a) = e$. Obviously, φ can be extended to a map $T_X \to X$ with the same property and $a = \delta^*(f, w)$ for some $w \in M^*$ is a subtree of f.

Conversely, let an endomorphism $\varphi : T_X \to T_X$ be given, mapping a subtree of f into E. Thus, $\varphi(\delta^*(f, w)) \in E$ for some $w \in M^*$. Assume $\mathcal{A} \not\models f$, then there is an $a \in A$ and a map $\psi : A \to X$ with $\tilde{\psi}(a) = f$. It follows with $\chi := \varphi \circ \tilde{\psi} : A \to T_X$ that $\chi(\delta^*(a, w)) = \varphi(\delta^*(\tilde{\psi}(a), w)) = \tilde{\varphi}(\delta^*(f, w)) \in E$, so $\mathcal{A} \not\models E$. $\square$

References

[GS98] H.P. Gumm and T. Schröder, *Covarieties and complete covarieties*, Coalgebraic Methods in Computer Science (B. Jacobs et al, ed.), Electronic Notes in Theoretical Computer Science, vol. 11, Elsevier Science, 1998, To appear in Theoretical Computer Science.

[GS99] H.P. Gumm and T. Schröder, *Products of coalgebras*, Algebra Universalis, to appear, 1999.

[GS00a] H.P. Gumm and T. Schröder, *Coalgebraic structure from weak limit preserving functors*, Coalgebraic Methods in Computer Science (H. Reichel, ed.), Electronic Notes in Theoretical Computer Science, vol. 33, Elsevier Science, 2000, pp. 113–133.

[GS00b] H.P. Gumm and T. Schröder, *Coalgebras of bounded type*, Tech. Report 25, FG Informatik, Philipps-Universität Marburg, 2000.

[Gum99a] H.P. Gumm, *Elements of the general theory of coalgebras*, LUATCS 99, Rand Afrikaans University, Johannesburg, South Africa, 1999.

[Gum99b] H.P. Gumm, *Equational and implicational classes of coalgebras*, Theoretical Computer Science, to appear, 1999.

[Rei95] H. Reichel, *An approach to object semantics based on terminal coalgebras*, Math. Struct. in Comp. Sci. (1995), no. 5, 129–152.

[Rut00] J.J.M.M. Rutten, *Universal coalgebra: a theory of systems*, Theoretical Computer Science **249** (2000), no. 1, 3–80.

[Trn69] V. Trnková, *Some properties of set functors*, Comm. Math. Univ. Carolinae **10** (1969), no. 2, 323–352.

Philipps-Universität Marburg, 35032 Marburg, Germany
E-mail address: `gumm@mathematik.uni-marburg.de`