Mathematical Notes

Eric G. Wagner – Wagner Mathematics

Notes on Birkhoff's Theorem[1]

Abstract

Sixty years ago Garrett Birkhoff, [1], published one of the major results in the area of universal algebra, namely the theorem that equational varieties (i.e., categories of algebras with equational presentations) are characterized by the following properties:

- they contain an algebra with a nonempty carrier.
- they are closed under subalgebras.
- they are closed under homomorphic images.
- they are closed under products.

These notes present a proof of the theorem, an analysis of the proof, a look at some of its consequences in mathematics and computer science. It also contains a brief review of other techniques for presenting algebras using universal formulae, namely conditional axioms (positive Horn formulae), general Horn formulae, and Gentzen formulae. We illustrate how the techniques gain in expressiveness at the cost of losing various closure properties. We also include a starting section that provides a review of basic algebraic concepts and results.

Contents

[1]These notes were originally prepared in October 1995 as part of the material for lectures at the University of Bergen on applications of universal algebra in computer science. This work has been partially supported by the Architectural Abstraction project under NFR (Norway), by CEC under ESPRIT-II Basic Research Working Group No. 6112 COMPASS and by the Department of Informatics at the University of Bergen.

1 Mathematical Preliminaries

1.1 Signatures, Algebras and Derived Operations

For any natural number $n \in \{0, 1, 2, \ldots\}$, let $[n] = \{1, 2, \ldots, n\}$. Then $[0] = \emptyset$ is the empty set.

Definition 1.1 Let S be a set, then a *string u on S of length n* is a mapping $u : [n] \to S$. Let S^n be the set of all strings on S of length n, let $S^* = \bigcup \langle S^n \mid n = 0, 1, 2, \ldots \rangle$. Let $\lambda : [0] \to S$ be the unique mapping from $\emptyset$ to S, call λ *the empty string.* If $u \in S^n$, and $i \in [n]$ then we often write u_i for $u(i)$, and we also write $u = u_1 \cdots u_n$. □

Given a string $u \in S^*$ we write $|u|$ to denote the length of u. Thus, for any string $u \in S^*$, $u : [|u|] \to S$.

Definition 1.2 Let S be a set (of *sorts*).
An *S-sorted signature* is an $(S^* \times S)$-indexed family of sets

$$\Sigma = \langle \Sigma_{w,s} \mid w \in S^*, s \in S \rangle.$$

An element $\sigma \in \Sigma_{w,s}$ is an *operator symbol* of *rank* $\langle w, s \rangle$, *arity* w, and *sort* s. □

Definition 1.3 For an S-sorted signature Σ, a *Σ-algebra* A consists of

- An S-indexed family of sets $U(A) = \langle A_s \mid s \in S \rangle$.
- A mapping $\sigma_A : A_{s_1} \times \ldots \times A_{s_n} \to A_s$ for each $\sigma \in \Sigma_{w,s}$ where $w = s_1 \cdots s_n$ in S^*, and $s \in S$.

The set A_s is the *carrier of sort s*, σ_A is the *operation* in A named by the operator σ, if $\sigma \in \Sigma_{\lambda,s}$ then σ_A is said to be a *constant* of A. □

For $w = s_1 \cdots s_n \in S^*$ we will often write A^w for $A_{s_1} \times \ldots \times A_{s_n}$ and $\overline{a} = \langle a_1, \ldots, a_n \rangle \in A^w$ for $\overline{a} = \langle a_1, \ldots, a_n \rangle \in A_{s_1} \times \ldots \times A_{s_n}$, that is, $a_1 \in A_{s_1}$, ..., $a_n \in A_{s_n}$. The product A^λ, sometimes called the *empty product*, is a one element set, $\{\langle\ \rangle\}$ – its sole element, $\langle\ \rangle$, being the empty tuple of elements of A. Note that the above definition says that, for $\sigma \in \Sigma_{\lambda,s}$, $\sigma_A : \{\langle\ \rangle\} \to A_s$ rather than that $\sigma_A \in A_s$, this distinction makes a difference in Proposition 1.11.

Definition 1.4 For Σ-algebras A and B, a *Σ-homomorphism* $h : A \to B$ is a triple

$$\langle A,\ \overline{h},\ B \rangle$$

where $\overline{h} = \langle h_s : A_s \to B_s \mid s \in S \rangle$ is an S-indexed set of mappings such that, for every $\sigma \in \Sigma_{w,s}$, with $w = w_1 \cdots w_n$, and for all $\langle a_1, \ldots, a_n \rangle \in A^w$

$$h_s(\sigma_A(a_1,\ldots,a_n)) = \sigma_B(h_{w_1}(a_1),\ldots,h_{w_n}(a_n)).$$

Denote the category with the Σ-algebras as objects and Σ-homomorphisms as morphisms by $\mathbf{Alg}_\Sigma$. □

When it is clear from context what we mean, we will normally write h for $\overline{h}$ and h_s.

Definition 1.5 Let S be a set. Define $\mathbf{Set}^S$ to be the category with, as objects, S-indexed sets $A = \langle A_s \mid s \in S\rangle$, and as morphisms $f : A \rightarrow B$, S-indexed families of mappings, $f = \langle f_s : A_s \rightarrow B_s \mid s \in S\rangle$. □

Definition 1.6 Define $U : \mathbf{Alg}_\Sigma \rightarrow \mathbf{Set}^S$ to be the functor that takes each Σ-algebra A to its carrier, $U(A) = \langle A_s \mid s \in S\rangle$ and each homomorphism $h : A \rightarrow B$ to $U(h) = \overline{h} = \langle h_s : A_s \rightarrow B_s \mid s \in S\rangle$. □

Proposition 1.7 *The functor $U : \mathbf{Alg}_\Sigma \rightarrow \mathbf{Set}^S$ has a left adjoint, $F : \mathbf{Set}^S \rightarrow \mathbf{Alg}_\Sigma$.*

That is, for each $X \in |\mathbf{Set}^S|$ there exists a Σ-algebra, denoted $\mathbf{T}_\Sigma(X)$, and $\eta_X : X \rightarrow U(\mathbf{T}_\Sigma(X))$ such that for every Σ-algebra A and $f : X \rightarrow U(A)$ there is a unique Σ-homomorphism $f^\sharp : \mathbf{T}_\Sigma(X) \rightarrow A$ such that $U(f^\sharp) \bullet \eta_X = f$.

For the special case where $X_s = \emptyset$ for all $s \in S$ we write $\mathbf{T}_\Sigma$ rather than $\mathbf{T}_\Sigma(X)$.

The construction of $\mathbf{T}_\Sigma(X)$ can be made quite concrete in a manner that provides a convenient notation for parts of our development. The carrier of the desired algebra is the set of all $\Sigma(X)$-expressions defined as follows.

Definition 1.8 Let $X = \langle X_s \mid s \in S\rangle$. We assume that $\Sigma_{w,s} \cap X_{s'} = \emptyset$ for all $w \in S^*$ and $s, s' \in S$.

Given an S-sorted signature Σ define the set of $\Sigma(X)$-expressions in the usual way as an S-ary set where:

If $x \in X_s$ then x is a $\Sigma(X)$-expression of sort s.[2]

If $\sigma \in \Sigma_{u,s}$ and for each $i = 1,\ldots,|u|$, t_i is a $\Sigma(X)$-expression of sort u_i then the string $\sigma(t_1,\ldots,t_{|u|})$ is a $\Sigma(X)$-expression of sort s.[3] Note, this includes the case $u = \lambda$ but for $\sigma \in \Sigma_{\lambda,s}$ we will generally write the corresponding term as 'σ' rather than '$\sigma(\ )$'.

We write $T_\Sigma(X)_s$ for the set of $\Sigma(X)$-expressions of sort s. Write $T_\Sigma(X)$ for the S-indexed set $T_\Sigma(X) = \langle T_\Sigma(X)_s \mid s \in S\rangle$. Define $\mathbf{T}_\Sigma(X)$ to be the Σ-algebra with carrier $U(\mathbf{T}_\Sigma(X)) = T_\Sigma(X)$ and such that $\sigma \in \Sigma_{u,s}$ if $t_i \in T_\Sigma(X)_{u_i}$ for $i = 1,\ldots,|u|$ then

$$\sigma_{\mathbf{T}_\Sigma(X)}(t_i,\ldots,t_{|u|}) = \sigma(t_1,\ldots,t_{|u|}) \in (T_\Sigma(X_w))_s.$$

The algebra $\mathbf{T}_\Sigma(X)$ is called *the free Σ-algebra generated by X.* □

Fact 1.9 $\mathbf{T}_\Sigma$ *is initial in* $\mathbf{Alg}_\Sigma$.

[2]This fixes the unit of the adjunction by forcing $\eta_X(x) = x \in U(\mathbf{T}_\Sigma(X))_s$ for every $s \in S$ and $x \in X_s$.

[3]For everything to work properly we need the precise treatment of strings introduced above, so this informal presentation should be used with care.

A case of particular interest for our applications is where X is "a set of variables sorted (typed) by a string w". Given $w \in S^*$ and a string $\overline{x}$ of length $|w|$ of disjoint S-sorted variable names, define X_w to be the S-indexed family of sets ("of variables") where, for each $s \in S$, $(X_w)_s = \{x_i \mid w_i = s\}$.

Definition 1.10 A $\Sigma(X_w)$-expression is often called a *derived operator*. Given a $\Sigma(X_w)$-expression t of sort s and a Σ-algebra A, we define the corresponding *derived operation*, $\tau_{t,A,w}$, in A to be the function

$$\begin{aligned}\tau_{t,A,w} : A^w &\to A_s \\ \overline{a} &\mapsto h^\sharp_{\overline{a}}(t),\end{aligned}$$

where $\overline{a} \in A_w$ and $h^\sharp_{\overline{a}}$ is defined, as in the remark following Proposition 1.7, from

$$\begin{aligned}h_{\overline{a}} : X_w &\to U(A) \\ x_i &\mapsto \overline{a}_i.\end{aligned}$$

□

Proposition 1.11 *Let A be a Σ-algebra, let X_w be a set of variables for $w \in S^*$, and let $\overline{a} = \langle a_1, \ldots, a_{|w|}\rangle \in A^w$. Then for $\Sigma(X_w)$-expressions*

$\tau_{x_i,A,w}(\overline{a}) = a_i$ *for* $i = 1, \ldots, |w|$ *and* $x_i \in (X_w)_{w_i}$.

$\tau_{\sigma(t_1,\ldots,t_n),A,w}(\overline{a}) = \sigma_A(\tau_{t_1,A,w}(\overline{a}), \ldots, \tau_{t_n,A,w}(\overline{a}))$ *for* $u \in S^*$, $n = |u|$ *and* $\sigma \in \Sigma_{u,s}$.

Proof

$\tau_{x_i,A,w}(\overline{a}) = h^\sharp_{\overline{a}}(x_i) = h_{\overline{a}}(x_i) = a_i$.

$\tau_{\sigma(t_1,\ldots,t_n),A,w}(\overline{a}) = h^\sharp_{\overline{a}}(\sigma(t_1, \ldots, t_n)) = \sigma_A(h^\sharp_{\overline{a}}(t_1), \ldots, h^\sharp_{\overline{a}}(t_n))$
$= \sigma_A(\tau_{t_1,A,w}(\overline{a}), \ldots, \tau_{t_n,A,w}(\overline{a}))$

□

Note that, for $\sigma \in \Sigma_{\lambda,s}$, we have $\tau_{\sigma,A,w} : A^w \to A_s = \sigma_A(\langle\ \rangle)$ and not $\tau_{\sigma,A,w} : A^w \to A_s = \sigma_A$ as it is frequently, but imprecisely, written.

1.2 Congruences

Definition 1.12 Let A be a Σ-algebra. Then a *congruence*, Q, on A is a family $Q = \langle Q_s \subseteq (A_s \times A_s) \mid s \in S\rangle$ of equivalence relations such that if $\sigma \in \Sigma_{w,s}$, and, for $i = 1, \ldots, |w|$, $\langle a_{i,1}, a_{i,2}\rangle \in Q_{w_i}$ then

$$\langle \sigma_A(a_{1,1}, \ldots, a_{1,n}), \sigma_A(a_{2,1}, \ldots, a_{2,n})\rangle \in Q_s.$$

For $a \in A_s$ let $[a]_Q$ denote its Q_s-equivalence class. We will drop the subscript when the congruence is clear from context. □

Definition 1.13 Given S-sorted signature Σ, a Σ-algebra A and congruence Q on A, define the *quotient algebra of A given by Q* to be the Σ-algebra, denoted A/Q, such that

$$(A/Q)_s = \{[a]_Q \mid a \in A_s\}$$

and such that for all $w \in S^*$, $s \in S$, $\sigma \in \Sigma_{w,s}$ and $[a_i]_Q \in (A/Q)_{w_i}$, $i = 1, \ldots |w|$,

$$\sigma_{A/Q}([a_1]_Q, \ldots [a_{|w|}]_Q) = [\sigma(a_1, \ldots, a_{|w|})]_Q.$$

$\square$

Definition 1.14 Let A be an S-sorted Σ-algebra. Let Q and Q' be congruences on A. We say *Q is smaller than Q'* (or *Q' is larger than Q*), and write $Q \subseteq Q'$, if for all $s \in S$, $Q_s \subseteq Q'_s$. $\square$

Definition 1.15 Let A be an S-sorted Σ-algebra. Let K be a set of congruences on A, Define $\bigcap K$ to be the intersection of elements of K, i.e., the S-indexed set $\bigcap K$ such that for each $s \in S$ the indexed set $(\bigcap K)_s = \bigcap_{Q \in K} Q_s$ consists of exactly those pairs $\langle a, a' \rangle \in A_s \times A_s$ such that $\langle a, a' \rangle \in Q$ for every $Q \in K$. $\square$

Fact 1.16 *Let A be an S-sorted Σ-algebra. Let K be a set of congruences on A. Then $\bigcap K$ is a congruence on A. In particular, if $K = \emptyset$ then, for every $s \in S$, $(\bigcap \emptyset)_s = A_s \times A_s$.*

Proposition 1.17 *Let $R = \langle R_s \subseteq A_s \times A_s \mid s \in S \rangle$ be an S-indexed family of relations on A. Then there exists a smallest congruence Q_R on A such that $R_s \subseteq Q_{R,s}$ for every $s \in S$.*

We call Q_R *the congruence generated by$R$.*

Proposition 1.18 *Let A be a Σ-algebra and let Q be a congruence on A. Define*

$$\begin{gathered} h_Q : A \rightarrow A/Q \\ h_s : a \mapsto [a]. \end{gathered}$$

Then h_Q is a Σ-homomorphism. It is also an epimorphism in the category $\mathbf{Alg}_\Sigma$.

Proposition 1.19 *Let A and B be Σ-algebras and let $h : A \rightarrow B$ be a Σ-homomorphism. Define*

$$\begin{gathered} Q_h = \langle Q_{h,s} \mid s \in S \rangle \\ Q_{h,s} = \{\langle a, a' \rangle \mid a, a' \in A_s,\ h(a) = h(a')\}. \end{gathered}$$

Then Q_h is a congruence on A.

Proposition 1.20 *Let A and B be Σ-algebras and let $h : A \rightarrow B$ be a Σ-homomorphism. Let Q be a congruence on A such that $Q \subseteq Q_h$. Then h factors uniquely through h_Q. That is, there exists a unique $k : A/Q \rightarrow B$ such that $k \bullet h_Q = h$.*

$$\begin{tikzcd} A \arrow[rr, "h"] \arrow[dr, "h_Q"'] & & B \\ & A/Q \arrow[ur, "k"'] & \end{tikzcd}$$

Proof Let k be such that for every $a \in A$, $k([a]_Q) = h(a)$. $\square$

We call Q_h, as defined in the preceding proposition, *the congruence induced by h.*

Corollary 1.21 *If A is a Σ-algebra and Q and Q' are congruences on A such that $Q \subseteq Q'$ then*

$$\begin{gathered} h_{Q,Q'} : A/Q \rightarrow A/Q' \\ (h_{Q,Q'})_s : [a]_Q \mapsto [a]_{Q'} \end{gathered}$$

is a Σ-homomorphism.

1.3 Equations and Conditional Axioms

Definition 1.22 A $\Sigma(X_w)$-*equation of sort* s is a pair $e = \langle L, R\rangle \in (T_\Sigma(X_w))_s \times (T_\Sigma(X_w))_s$. We often write $L = R$ rather than $\langle L, R\rangle$.

A *conditional* $\Sigma(X_w)$-*axiom with* p *conditions (premises)* is an $\{1, \ldots, p, p+1\}$-indexed set of Σ-equations, $\langle e_i = \langle L_i, R_i\rangle \mid i = 1, \ldots, p+1\rangle$. We frequently write this as

$$e_1 \ \& \ \ldots \ \& \ e_p \Rightarrow e_{p+1}.$$

or as

$$(L_1 = R_1)\& \ldots \&(L_p = R_p) \Rightarrow (L_{p+1} = R_{p+1}).$$

□

Note that, a conditional axiom without any conditions, i.e., with $p = 0$, is an equation.

Definition 1.23 A $\Sigma(X_w)$-equation $e = \langle L, R\rangle$ Σ-*holds* in a Σ-algebra A if, and only if, for every $\overline{a} \in A^w$, we have $\tau_{L,A,w}(\overline{a}) = \tau_{R,A,w}(\overline{a})$.

A $\Sigma(X_w)$-conditional axiom

$$e_1 \ \& \ \ldots \ \& \ e_p \Rightarrow e_{p+1}$$

where $e_i = \langle L_i, R_i\rangle$, $i = 1, \ldots, p+1$, Σ-*holds* in a Σ-algebra A if, and only if, for every $\overline{a} \in A^w$, if $\tau_{L_i,A,w}(\overline{a}) = \tau_{R_i,A,w}(\overline{a})$ for $i = 1, \ldots, p$ then $\tau_{L_{p+1},A,w}(\overline{a}) = \tau_{R_{p+1},A,w}(\overline{a})$. □

Definition 1.24 Let C be a set of conditional axioms. Any Σ-algebra that Σ-holds for every conditional axiom in C is said to Σ-*satisfy* C, and is called a $\langle\Sigma, C\rangle$-*algebra*. The category with $\langle\Sigma, C\rangle$-algebras and Σ-homomorphisms between them is denoted $\mathbf{Alg}_{\Sigma,C}$. □

The set C is often called a Σ-*specification* or a Σ-*presentation*. The categories $\mathbf{Alg}_{\Sigma,C}$ are often called *quasi-varieties* – see Cohn, [2].

Fact 1.25 *For every set C of conditional axioms, $\mathbf{Alg}_{\Sigma,C}$ is a full subcategory of $\mathbf{Alg}_\Sigma$.*

Definition 1.26 Let $c = [e_1 \ \& \ \ldots \ \& \ e_p \Rightarrow e_p]$ be a $\Sigma(X_w)$-conditional axiom with $e_i = \langle L_i, R_i\rangle$ for $i = 1, \ldots, |w|$. Let A be a Σ-algebra and let Q be a congruence on A. Then Q Σ-*satisfies* c iff for every $\overline{a} \in A^w$, if $\langle\tau_{L_i,A,w}(\overline{a}),\ \tau_{R_i,A,w}(\overline{a})\rangle \in Q$ for $i = 1, \ldots, p$ then $\langle\tau_{L_{p+1},A,w}(\overline{a}),\ \tau_{R_{p+1},A,w}(\overline{a})\rangle \in Q$. □

Proposition 1.27 *Let C be a set of Σ-conditional axioms. Then the initial $\langle\Sigma, C\rangle$-algebra is $I_C =_{def} \mathbf{T}_\Sigma/\mathbf{Q}$ where $\mathbf{Q}$ is the smallest congruence on $\mathbf{T}_\Sigma$ Σ-satisfying C.*

Proof Let $\mathcal{Q}$ be the set of all congruences on $\mathbf{T}_\Sigma$ that Σ-satisfy C. $\mathcal{Q} \neq \emptyset$ since $\mathcal{Q}$ contains the congruence max such that for every $s \in S$, $max_s = (T_\Sigma)_s \times (T_\Sigma)_s$. Let $\mathbf{Q} = \bigcap \mathcal{Q}$. It is easily shown that $\mathbf{Q}$ is a congruence Σ-satisfying C. Thus, $I_C =_{def} \mathbf{T}_\Sigma/\mathbf{Q}$ is well-defined.

To see that I_C is initial, let A be any $\langle\Sigma, C\rangle$-algebra, then by the initiality of $\mathbf{T}_\Sigma$ there exists a unique homomorphism $h : \mathbf{T}_\Sigma \to A$. Define Q_h on $\mathbf{T}_\Sigma$ as in Proposition 1.19, then Q_h Σ-satisfies C, hence $\mathbf{Q} \subseteq Q_h \in \mathcal{Q}$. Then by Proposition 1.20, h factors uniquely through $\mathbf{T}_\Sigma/\mathbf{Q} = I_C$. □

2 Closure Properties

It is generally useful to know if a subcategory $\mathcal{K} \subseteq \mathbf{Alg}_\Sigma$ is closed under certain constructions such as product, subalgebras, homomorphic images, coproducts, etc. In this section we review the definitions of a number of such closure properties for one-sorted algebras. These closure properties can, not surprisingly, be generalized to many-sorted algebras but some of the definitions become a bit messier.

There are two properties a subcategory $\mathcal{K} \subseteq \mathbf{Alg}_\Sigma$ may have that may be regarded as "essential" in that if $\mathcal{K}$ lacks these properties then we are not going to have very many tools to work with or we are at least going to have to take great care that we do not use them incorrectly. These basic properties are

abstractness A subcastegory $\mathcal{K} \subseteq \mathbf{Alg}_\Sigma$ is said to be *abstract* if whenever $A \in |\mathcal{K}|$ and $\alpha : A \to B$ is an isomorphism in $\mathbf{Alg}_\Sigma$ then both B and h are in $\mathcal{K}$. This property could also be described as "closure under isomorphisms", however that phrase is open to other interpretations.

regularity A subcategory $\mathcal{K} \subseteq \mathbf{Alg}_\Sigma$ is said to be *regular* if every homomorphism $h : A \to B$ in $\mathcal{K}$ is such that where Q_h is the congruence generated by h we have that A/Q_h, and the homomorphisms

$$\begin{gathered} h_{Q_h} : A \to A/Q_h \\ a \to [a]_{Q_h} \end{gathered}$$

and

$$\begin{gathered} k_h : A/Q_h \to B \\ [a] \mapsto h(a) \end{gathered}$$

are in $\mathcal{K}$.

Definition 2.1 A category $\mathcal{K}$ of Σ-algebras is said to be *closed under subalgebras* if for every $A \in \mathcal{K}$, if B is a subalgebra of A via monomorphism $m : B \to A$ in $\mathbf{Alg}_\Sigma$ then both B and m are in $\mathcal{K}$. □

Definition 2.2 A category $\mathcal{K}$ of Σ-algebras is said to be *closed under homomorphic images* if for every $A \in \mathcal{K}$, $B \in \mathbf{Alg}_\Sigma$ and $h : A \to B$ in $\mathbf{Alg}_\Sigma$, we have that $Im(h) \in \mathcal{K}$ so is the corestriction of h to $Im(h)$, i.e., homomorphism, denoted, $h \rceil Im(h) : A \to Im(h)$, that we get form h cutting back its target to $Im(h)$. □

Definition 2.3 A category $\mathcal{K}$ of Σ-algebras is said to be *closed under products* if for every set I and I-indexed family $\langle A_i \mid i \in I \rangle$ of algebras from $\mathcal{K}$ their product $\prod A_i =_{def} \prod \langle A_i \mid i \in I \rangle$ is also in $\mathcal{K}$ along with the associated projections $\pi_i : \prod A_j \to A_i$ and mediating morphisms (i.e., if $\langle f_i : B \to A_i \mid i \in I \rangle$ is an I-indexed family of morphisms in $\mathcal{K}$ then and $f : B \to \prod A_i$ in $\mathbf{Alg}_\Sigma$ such that $\pi_i \bullet f = f_i$ for every $i \in I$ then f is in $\mathcal{K}$.

Sometimes it is useful to break closure under products into two cases:

- Closure under empty products – that is, that $\mathcal{K}$ contains an algebra with exactly one element in its carrier. We leave it as an exercise for the reader to show that the product of an empty set of algebras is such a "one point" algebra.

- Closure under products of non-empty sets of algebras. The product of a set of algebras that contains just one element, A, is just some algebra isomorphic to A along with the isomorphism between them (another exercise for the reader). Thus products of non-empty sets of algebras are not all that interesting unless there are at least two algebras.

$\square$

We will make use of the following:

Proposition 2.4 *If a category of Σ-algebras, $\mathcal{K}$, is closed under homomorphic images then it is abstract, regular, and, for every $h : A \to B$ where $A \in \mathcal{K}$ and $B \in \mathbf{Alg}_\Sigma$, we have $Im(h) \cong A/Q_h$ where Q_h is the congruence induced by h – see Proposition 1.20 and remark following it.*

Proof If $A \in \mathcal{K}$ and $B \in \mathbf{Alg}_\Sigma$ such that $A \cong B$, then there must exist an isomorphism $\iota : A \to B$. But clearly $B = Im(h)$ and $\iota = \iota\lceil B$, so B and ι are in $\mathbf{K}$ by closure under homomorphic images. Thus $\mathcal{K}$ is abstract.

Clearly $A/Q_h \cong Im(h)$ from which the rest follows directly. $\square$

3 The Birkhoff Theorem

For the sake of simplicity we restrict out attention to one-sorted algebras.

Theorem 3.1 *(The Birkhoff Theorem) Let Σ be a signature, and let $\mathcal{K}$ be a full subcategory of $\mathbf{Alg}_\Sigma$. Then $\mathcal{K}$ is a variety iff the following four conditions hold for $\mathcal{K}$*

B.1. $\mathcal{K}$ contains an algebra with a non-empty carrier. (Note this condition is redundant if we interpret B.4, below, as including empty products).

B.2. $\mathcal{K}$ is closed under subalgebras.

B.3. $\mathcal{K}$ is closed under homomorphic images. (recall that this implies that $\mathcal{K}$ is abstract and regular.

B.4. $\mathcal{K}$ is closed under products.

Proof We give the proof subject to the assumption that, for any sufficiently large set X, that $\mathcal{K}$ contains an algebra $F(X)$ freely generated by X w.r.t. $\mathcal{K}$, that is, $F(X)$ is equipped with an injective mapping $\nu_X : X \to U(F(X))$ such that, for any $\mathcal{K}$-algebra A and mapping $f : X \to U(A)$ there exists a unique homomorphism $\overline{f} : F(X) \to A$ (in $\mathcal{K}$ since it is full) such that $I(\overline{f}) \bullet \nu_X = f$. We subsequently show that the existence of such free algebras is a consequence of the Birkhoff conditions.

It is easy to show that any (one-sorted) variety satisfies the Birkhoff conditions. Note, in particular, that any variety contains the algebra with but one element in its carrier.

It is also easy to see that there is a trivial variety, with but one algebra, namely the variety given by the equation $[x = y]$. Henceforth we will assume that $\mathcal{K}$ is non-trivial (contains an algebra with more than one element in its carrier).

We proceed as follows:

Given $\mathcal{K}$ let E be the set of all equations that are satisfied by all the algebras in $\mathcal{K}$. Then $\mathbf{Alg}_{\Sigma,E}$ is a variety and clearly $\mathcal{K} \subseteq \mathbf{Alg}_{\Sigma,E}$, so what we want to do is establish the equality.

Let $A \in \mathbf{Alg}_{\Sigma,E}$ and let X be a generating set for A with inclusion mapping $g : X \to U(A)$. Let $W(X)$ be the $\Sigma(X)$-word algebra (where $\Sigma(X)$ is the signature resulting from adding the elements of X to Σ_0) and let $\eta_X : X \to U(W(X))$ be the corresponding unit of the associated adjunction. Then there exists a unique homomorphism $g^\sharp : W(X) \to A$ in $\mathbf{Alg}_{\Sigma,E}$ such that $U(g^\sharp) \bullet \eta_X = g$.

$$\begin{tikzcd} X \arrow[r, "\eta_X"] \arrow[dr, "g"'] & U(W(X)) \arrow[d, "U(g^\sharp)"] & W(X) \arrow[d, "g^\sharp"] \\ & U(A) & A \end{tikzcd}$$

Furthermore, $g^\sharp$ is a surjective epimorphism since it hits all the generators of A.

Let $F(X)$, equipped with $\nu_X : X \to U(F(X))$, be the free algebra generated by X w.r.t. $\mathcal{K}$ that we are assuming exists. Then there exists a unique homomorphism $\nu_X^\sharp : W(X) \to F(X)$ in $\mathbf{Alg}_{\Sigma,E}$ such that $U(\nu_X^\sharp) \bullet \eta_X = \nu_X$.

$$\begin{tikzcd} X \arrow[r, "\eta_X"] \arrow[dr, "\nu_X"'] & U(W(X)) \arrow[d, "U(\nu_X^\sharp)"] & W(X) \arrow[d, "\nu_X^\sharp"] \\ & U(F(X)) & F(X) \end{tikzcd}$$

Again $\nu_X^\sharp$ is a surjective epimorphism because it hits all the generators of $F(X)$. Let $Q_{\nu_X^\sharp}$ be the congruence on $W(X)$ induced by $\nu_X^\sharp$.

By definition, $\langle t, t' \rangle \in Q_{\nu_X^\sharp}$ iff $\nu_X^\sharp(t) = \nu_X^\sharp(t')$.

Thus if $\langle t, t' \rangle \in Q_{\nu_X^\sharp}$ then the equation $t = t'$ holds in $F(X)$. But then this equation holds in all algebras $K \in \mathcal{K}$ for given an assignment $\alpha : X \to U(K)$ we get homomorphisms

$$\alpha^\sharp : W(X) \to K$$

$$\overline{\alpha} : F(X) \to K$$

But $\alpha^\sharp = \overline{\alpha} \bullet \nu_X^\sharp$ since $\overline{\alpha} \bullet \nu_X^\sharp \bullet \eta_X = \overline{\alpha} \bullet \nu_X = \alpha$ and because of the universality of $W(X)$.

$$
\begin{tikzcd}
X \arrow[r, "\eta_X"] \arrow[dr, "\nu_X"] \arrow[ddr, "\alpha"'] & U(W(X)) \arrow[d, "U(\nu^\sharp_X)"] \arrow[dd, bend left=60, "U(\alpha^\sharp)"] \\
& U(F(X)) \arrow[d, "U(\overline{\alpha})"] \\
& U(K)
\end{tikzcd}
\qquad
\begin{tikzcd}
F(X) \arrow[dd, "\overline{\alpha}"] \\
\\
K
\end{tikzcd}
\qquad
\begin{tikzcd}
W(X) \arrow[dd, "\alpha^\sharp"] \\
\\
K
\end{tikzcd}
$$

Now, by definition, $\langle t, t' \rangle$ holds in K at α iff $\alpha^\sharp(t) = \alpha^\sharp(t')$. But, for $\langle t, t' \rangle \in Q_{\nu^\sharp_X}$ we have $\alpha^\sharp(t) = \overline{\alpha} \bullet \nu^\sharp_X(t) = \overline{\alpha} \bullet \nu^\sharp_X(t') = \alpha^\sharp(t')$ since $\nu^\sharp_X(t) = \nu^\sharp_X(t')$ so $\langle t, t' \rangle$ is satisfied in K at α. So, in particular, A satisfies $\langle t, t' \rangle$ and so $\langle t, t' \rangle \in Q_{g^\sharp}$.

But then, $Q_{\nu^\sharp_X} \subseteq Q_{g^\sharp}$ and so, by Proposition 1 .20, $g^\sharp$ factors through $\nu^\sharp_X : W(X) \to F(X) = W(X)/Q_{\nu^\sharp_X}$ via a unique homomorphism g^*.

$$
\begin{tikzcd}
W(X) \arrow[rr, "g^\sharp"] \arrow[dr, "\nu^\sharp_X"'] & & A \\
& F(X) \arrow[ur, "g^*"'] &
\end{tikzcd}
$$

But since $g^\sharp$ is a surjective epimorphism it follows that g^* is a surjective epimorphism, and thus that A is a homomorphic image of $F(X)$. But then, by Birkhoff condition B.3, A is in $\mathcal{K}$ and we are done! $\square$

Proposition 3.2 *Let $\mathcal{K}$ be a subcategory of $\mathbf{Alg}_\Sigma$ that satisfies the Birkhoff conditions and is non-trivial (i.e., it contains an algebra with more than one element in its carrier). Then $\mathcal{K}$ has a "universal functor", that is, for every set X there exists an algebra $V(X) \in \mathcal{K}$ and a set mapping $\nu_X : X \to U(V(X))$ such that, for any algebra $B \in \mathcal{K}$ and mapping $f : X \to U(B)$ there is a unique morphism $\overline{f} \in \mathcal{K}(V(X), B)$ such that $U(\overline{f}) \bullet \nu_X = f$.*

Proof Let X be given and let $W = W(X)$ be the Σ-algebra freely generated by X and let $\eta_X : X \to U(W(X))$ be the associated universal arrow.

Define $\mathcal{Q}$ to be the set of all congruences Q on W such that $W/Q \in \mathcal{K}$. Then, by the Birkhoff condition B.4,

$$A =_{def}= \prod \langle W/Q \mid Q \in \mathcal{Q} \rangle \in \mathcal{K},$$

with projections $\pi_Q : A \to W/Q$ in $\mathcal{K}$.

For each $Q \in \mathcal{Q}$ let $h_Q : W \to W/Q$ be the homomorphism induced by Q. Then for each $Q \in \mathcal{Q}$, $U(h_Q) \bullet \eta_X : X \to U(W/Q)$ and so, since U preserves products, there exists a unique $\sigma_X : X \to U(A)$ such that $U(\pi_Q) \bullet \sigma_X = U(h_Q)$ for each $q \in \mathcal{Q}$.

Let $V(X)$ be the subalgebra of A generated by the image of σ_X.

Let $i : V(X) \to A$ be in inclusion homomorphism given by $V(X)$ being a subalgebra of A.

By Birkhoff condition B.2 we have $V(X) \in \mathcal{K}$ and $i \in \mathcal{K}$.

Let ν_X be the corestriction of σ_X to $U(V(X))$. We claim that $\langle V(X), \nu_X \rangle$ is the desired universal arrow.

The argument goes as follows:

$$\sigma_X = U(i) \bullet \nu_X \tag{1}$$

we also have

$$U(\pi_Q) \bullet \sigma_A = U(h_Q) \bullet \eta_A \tag{2}$$

for all $Q \in \mathcal{Q}$.

Now let B be an arbitrary $\mathcal{K}$-algebra and let $\alpha : X \to U(B)$. Without loss of generality we may assume that B is generated by the image of α. By the freeness of $W = W(X)$ with respect to X, α extends uniquely to a homomorphism $\beta : W \to B$ such that

$$U(\beta) \bullet \eta_X = \alpha \tag{3}$$

Now β is an epimorphism because it hits all the generators of B. Let Q_β be the congruence on W induced by β and let $h_{Q_\beta} : W \to W/Q_\beta$ be the indicated homomorphism induced by Q_β. Then by Proposition 1.20 there exists a unique homomorphism $\beta^* : W/Q_\beta \to B$ such that

$$\beta = \beta^* \bullet h_{Q_\beta} \tag{4}$$

But β epic implies β^* epic and so β^* is an isomorphism since β^*, by construction, is a monomorphism. That is, by the definition in the proof of the proposition, $\beta^*([a]) = \beta(a)$ so $\beta^*([a]) = \beta^*([b])$ would imply $\beta(a) = \beta(b)$ which, by the definition of Q_β, implies $[a] = [b]$.

But if β^* is an isomorphism then $W/Q_\beta \in \mathcal{K}$, since $\mathcal{K}$ is abstract, and so $Q_\beta \in \mathcal{Q}$. Then, letting $\overline{\alpha}$ be the homomorphism

$$\overline{\alpha} = \beta^* \bullet \pi_{Q_\beta} \bullet i \tag{5}$$

we get the diagram

$$\begin{tikzcd}
X \arrow[rr, "\nu_X"] \arrow[dr, "\alpha"] \arrow[dd, "\eta_X"'] & & U(V(X)) \arrow[dl, "U(\overline{\alpha})"'] \arrow[dd, "U(\pi_{Q_\beta} \bullet i)"] \\
& U(B) & \\
U(W) \arrow[ur, "U(\beta)"] \arrow[rr, "U(h_{Q_\beta})"'] & & U(W/Q_\beta) \arrow[ul, "U(\beta^*)"']
\end{tikzcd}$$

where the square, and all but possible the top, triangle commute

$\triangleright$ by (3)

$\triangle$ by (4)

$\triangleleft$ by inspection.

$\square$ as follows:

$$\begin{array}{lll} U(\pi_{Q_\beta} \bullet i) \bullet \nu_X & & \\ = & U(\pi_{Q_\beta}) \bullet U(i) \bullet \nu_X & U \text{ a functor} \\ = & U(\pi_{Q_\beta}) \bullet \sigma_X & (1) \\ = & U(h_{Q_\beta}) \bullet \eta_X & (2) \end{array}$$

whence, by diagram chasing

$$\begin{array}{lll} \alpha & & \\ = & U(\beta) \bullet \eta_X & \triangleright \\ = & U(\beta^*) \bullet U(h_{Q_\beta}) \bullet \eta_A & \triangle \\ = & U(\beta^*) \bullet U(\pi_{Q_\beta} \bullet i) \bullet \nu_X & \square \\ = & U(\overline{\alpha}) \bullet \nu_X & \text{def. } \overline{\alpha} \end{array}$$

so ∇ commutes as required. Finally, $\overline{\alpha}$ is unique for if $U(\alpha'') \bullet \nu_X = \alpha = U(\overline{\alpha}) \bullet \nu_X$ then α'' and $\overline{\alpha}$ agree on the image of ν_X and so are equal. $\square$

Note that in the above proof that $V(X)$ is generated by the image of $\nu_X : X \to U(V(X))$. We say that $V(X)$ is a *free-algebra generated by* X if ν_X is injective.

Proposition 3.3 *If $\mathcal{K}$ is a nontrivial category of Σ-algebras and there is a universal functor for U, then there is a cardinal number α such that for every set X with $|X| \geq \alpha$ there is a free algebra $V(X)$ generated by X in $\mathcal{K}$. Furthermore, every algebra A in $\mathcal{K}$ is the homomorphic image of a free algebra in $\mathcal{K}$.*

Proof By hypothesis there exists an algebra A in $\mathcal{K}$ with more than one element in its carrier. Let X_0 be a set of generators for A and let $\nu_{X_0} : X_0 \to U(V(X_0))$ be the universal mapping given by the preceding proposition. Then, where $g : X_0 \to U(A)$ is the evident inclusion mapping (and an injection), we get that there exists a unique homomorphism $\overline{g} : V(X_0) \to A$ such that $U(\overline{g}) \bullet \nu_{X_0} = g$.

$$\begin{tikzcd} X_0 \arrow[r, "\nu_{X_0}"] \arrow[dr, "g"'] & U(F(X_0)) \arrow[d, "U(\overline{g})"] & F(X_0) \arrow[d, "\overline{g}"] \\ & U(A) & A \end{tikzcd}$$

But then g injective forces ν_{X_0} injective.

We claim $\alpha = |X_0|$. Let Y be any set of cardinality at least $|X_0|$ and let $\nu_Y : Y \to U(V(Y))$ be the correspondinguniversal arrow given by the preceding proposition. Now let $y, y' \in Y$, $y \neq y$,, we need to show that $\nu_Y(y) \neq \nu_Y(y')$. But clearly we can find a surjection $\rho : Y \to X_0$ such that $\rho(y) \neq \rho(y')$, and then we have $U(\overline{\nu_{X_0} \bullet \rho}) \bullet \nu_Y = \nu_{X_0} \bullet \rho$ but $\nu_{X_0}(\rho(y)) \neq \nu_{X_0}(\rho(y'))$ so we must have $\nu_Y(y) \neq \nu_Y(y')$, thus ν_Y must be injective. $\square$

4 Discussion of the proofs

4.1 The proof of the Birkhoff theorem

The theorem itself – assuming $F(X)$ exists:

Given that $\mathcal{K}$ satisfies the Birkhoff conditions and $F(X)$ exists we want to show that it is a variety, i.e., that there exists a set of equations, E, such that $\mathcal{K} = \mathbf{Alg}_{\Sigma,E}$.

Proposal: Take E to be the set of all equations satisfied by all the algebras in $\mathcal{K}$.

Observation: If $A \in \mathbf{Alg}_{\Sigma,E}$ and for some X such that $F(X)$ is in $\mathcal{K}$, we can show that there is a surjective homomorphism $F(X) \rightarrow A$ then we will have $A \in \mathcal{K}$ by closure under homomorphic images.

Observation: If $A \in \mathcal{K}$, X is a set of generators for A, $g : X \rightarrow U(A)$ is the inclusion mapping, then $g^\sharp : W(X) \rightarrow A$ is surjective.

Observation: If $\alpha \bullet \beta = h$, a surjective homomorphism, then α is surjective.

Proposal: Show that $g^\sharp$ factors through $F(X)$.

Observation: We have $\nu_X^\sharp : W(X) \rightarrow F(X)$

Proposal: Try to factor through ν_X.

Observation: If $Q_{\nu_X^\sharp} \subseteq Q_{g^\sharp}$ then a factoring through $\nu_X^\sharp$ will exist.

Proposal: Show that $Q_{\nu_X^\sharp} \subseteq Q_{g^\sharp}$.

Observation: If $A \in \mathcal{K}$, X is a set of generators for A, $g : X \rightarrow U(A)$ is the inclusion mapping, and $\langle t, t' \rangle$ is a $\Sigma(X)$-equation in E, then $g^\sharp(t) = g^\sharp(t')$ so $\langle t, t' \rangle \in Q_{g^\sharp}$.

Proposal: Show that $Q_{\nu_X^\sharp} \subseteq E$. I.e., that for every $K \in \mathcal{K}$, $\alpha : X \rightarrow U(K)$, and $\langle t, t' \rangle \in Q_{\nu_X^\sharp}$ we have $\alpha^\sharp(t) = \alpha^\sharp(t')$.

Observation: By definition, $\nu_X^\sharp(t) = \nu_X^\sharp(t')$ for every $\langle t, t' \rangle \in Q_{\nu_X^\sharp}$.

Proposal: Show that $\alpha^\sharp$ factors through $\nu_X^\sharp$.

Observation: Easy consequence of universality.

4.2 The Universal Functors theorem

We want to show that if $\mathcal{K}$ is abstract, regular and closed under subalgebras and products that it has a universal functor. That is, given X we have $\langle V(X),\ \nu_X : X \rightarrow U(V(X)) \rangle$ such that:

1. For every $A \in \mathcal{K}$ and $f : X \to U(A)$ there exists a unique $\overline{f} : V(X) \to A$ such that $U(\overline{f}) \bullet \nu_A = f$,
2. $V(X) \in \mathcal{K}$.

Observation: We have a universal functor $\langle W(X),\ \eta_A : X \to U(W(X)) \rangle \in \mathbf{Alg}_\Sigma$. So $W(X)$ satisfies (1) but generally does not satisfy (2).

Proposal: Perhaps we can define $V(X)$ as a quotient of $W(X)$.

Observation: For any algebra $A \in \mathcal{K}$ and homomorphism $h : W(X) \to A$ we have, by the regularity of $\mathcal{K}$, that h factors as $h^* \bullet h_{Q_h}$ through $W(X)/Q_h$, where h^* is a monomorphism and so, $W(X)/Q_h$ is a subalgebra of A (up to an isomorphism supplied by the abstractness of $\mathcal{K}$).

Observation: That means each $W(X)/Q_h \in \mathcal{K}$ by the closure of $\mathcal{K}$ under subalgebras.

Proposal: Find an $\mathcal{A} \in \mathcal{K}$ and an $\sigma : W(X) \to \mathcal{A}$ such that $V(X)$ is $W(X)/Q_\sigma$.

Observation: $V(X)$ must be such that each $h_Q : W(X) \to W(X)/Q$ factors through $V(X)$.

Proposal: Combine all these $W(X)/Q$ in some way to get a $\mathcal{K}$-algebra $\mathcal{A}$ and a $\sigma : W(X) \to \mathcal{A}$ through which we can factor all the h_Q then $W/Q_h \in \mathcal{K}$,

Observation: $\mathcal{Q} = \{Q \mid W(X)/Q \in \mathcal{K}\}$ is a set (even though $|\mathcal{K}|$ is a most likely a proper class) and thus $\prod W(X)/Q \mid Q \in \mathcal{Q}$ exists and, by the Birkhoff Theorem, is in $\mathcal{K}$. Furthermore, there exists a unique morphism $\sigma : W(X) \to \prod W(X)/Q \mid Q \in \mathcal{Q}$ such that $\pi_Q \bullet \sigma = h_Q$ for every $Q \in \mathcal{Q}$.

Proposal: Try this choice of σ. That is, try $V(X) =_{def} im(\sigma) \cong W(X)/Q_\sigma$.

The remainder of the proof is really a fairly straightforward matter of showing that the above idea works. Note that what the use of the product really does is give us a way to pick an element of $\{W(X)/Q \mid Q \in \mathcal{Q}\}$ which is "less-than" all the others if we look at the look at the existence of a morphism $g : A \to B$ as a kind of "less-than" relation.

4.3 Another example of the same proof idea

Proposition 4.1 *Let Σ be a signature and let E be a set of $\Sigma(X)$-equations. Then there exists a Σ-algebra A/E and an epimorphism $h_E : A \to A/E$ such that*

- *A/E satisfies the axioms in E.*
- *A/E is a quotient algebra of A.*
- *A/E is such that if $h : A \to B$ where B satisfies E then there exists a unique $\overline{h} : A/E \to B$ such that h factors through $\overline{h}$.*

Proof If $h : A \to B$ where B satisfies E then h factors as $m_h \bullet e_h$ where $e_h : A \to Q/Q_h$ and m_h is injective. But then, by the Birkhoff theorem, A/Q satisfies E. Let $\mathcal{Q} = \{Q \mid A/Q \text{ satisfies } E\}$. Then $\mathcal{Q}$ is a set, and so we can form

$$\mathcal{A} = \prod \langle A/Q \mid Q \in \mathcal{Q} \rangle.$$

By the Birkhoff theorem, $\mathcal{A}$ satisfies E. For each $Q \in \mathcal{Q}$ let $\pi_Q : \mathcal{A} \to A/Q$ be the projection from $\mathcal{A}$ to A/Q. For each $Q \in \mathcal{Q}$ we have $h_Q : A \to A/Q$ ($h_Q : a \mapsto [a[_Q)$. By the definition of products there exists a unique $g : A \to \mathcal{A}$ such that $\pi_Q \bullet g = h_Q$ for each $Q \in \mathcal{Q}$.

Claim: what we want is $A/E =_{def} A/Q_g$.

The argument goes as follows: Let $h : A \to B$ where B satisfies E. Then, as above, we get $Q_h \in \mathcal{Q}$ and there there exist $h^* : A/Q_h \to B$. But, by the above $h_{Q_h} = \pi_{Q_h} \bullet g^* \bullet h_{Q_g}$. So $h = h^* \bullet \pi_{Q_h} \bullet g^* \bullet h_{Q_g}$. That is, take $\overline{h} = h^* \bullet \pi_{Q_h} \bullet g^*$. The uniqueness follows from h_{Q_g} being an epimorphism. □

5 Some Simple Applications of the Birkhoff Theorem

We give a selection of examples of the use of the Birkhoff theorem to show that certain categories of Σ-algebras are not varieties. Of course the theorem can also be used to show that particular categories of Σ-algebras are varieties, but we shall not give any examples of that here.

5.1 One-to-One Algebras

Consider the following signature, One:

sorts: $\{\mathsf{Data}, \mathsf{Labels}\}$

operators:

mark:Data→Labels

Now let $\mathcal{K}$ be the category of One-algebras, A, such that $\mathsf{mark}_A : A_{\mathsf{Data}} \to A_{\mathsf{Labels}}$ is injective (i.e., a a one-to-one function).

Question: Is $\mathcal{K}$ a variety.

Answer: No, because it is not closed under homomorphic images. Specifically, let A be a $\mathcal{K}$-algebra where $|A_{\mathsf{Labels}}|$ has more than one element, then given $a, b \in A_{\mathsf{Labels}}$ with $a \neq b$, it is easy to see that the set $\{\langle a, b \rangle\}$ generates a congruence, Q, on A that that identifies a and b but does not identify and other elements of A_{Labels} nor any elements of A_{Data}, then the homomorphism $h : A \to A/Q$ has A/Q as its homomorphic image but $A/Q \notin \mathcal{K}$ because $\mathsf{mark}_{A/Q}(\mathsf{mark}_A^{-1}(a)) = \mathsf{mark}_{A/Q}(\mathsf{mark}_A^{-1}(b))$ but $\mathsf{mark}_A^{-1}(a) \neq \mathsf{mark}_A^{-1}(b)$ in A/Q.

5.2 The Nice-Bool algebras

Intuitively the data type Bool has exactly two elements, true and false, and in any situation where our signature, Σ, contains a sort Bool we would like to restrict ourselves to algebras that satisfy this intuition. But it is easy to see that this the category $\mathcal{K}$ of Σ-algebras, A with $A_{\mathsf{Bool}} = \{\mathsf{true}_A, \mathsf{false}_A\}$ with $\mathsf{true}_A \neq \mathsf{false}_A$ is not a variety.

- If $A \in \mathcal{K}$ then $A^0 \notin \mathcal{K}$ because $|A^0_{\mathsf{Bool}}| = 1$.
- If $A \in \mathcal{K}$ then $A \times A \notin \mathcal{K}$ because $|(A \times A)_{\mathsf{Bool}}| = 4$.

5.3 If-Then-Else Algebras

Consider the following signature, IF:

sorts: $\{\mathsf{Data}\}$.

operators:

$\mathsf{if}{:}\mathsf{Data}\times\mathsf{Data}\times\mathsf{Data}\times\mathsf{Data}\to\mathsf{Data}$

Now let $\mathcal{K}$ be the category of all "IF-algebras", A where if_A behaves as we expect it to, that is

If A is such that for all $d, d', d'', d''' \in A_{\mathsf{data}}$ we have

$$\mathsf{if}_A(d, d', d'', d''') = \begin{cases} d'' & \text{if } d = d' \\ d''' & \text{if } d \neq d' \end{cases}$$

Such algebras clearly exist – e.g., take A to be such that

$A_{\mathsf{Data}} = \mathbf{Z}$ (the integers)

$$\mathsf{if}_A(z_1, z_2, z_3, z_4) = \begin{cases} z_3 & \text{if } z_1 = z_2 \\ z_4 & \text{if } z_1 \neq z_2 \end{cases}$$

Question: Can we define $\mathcal{K}$ using equations?

Answer: No, we can not because $\mathcal{K}$ is not closed under products. Specifically, letting A be the $\mathcal{K}$-algebra given immediately above, it is easy to see that $A \times A$ is not in $\mathcal{K}$ because we have

$$\mathsf{if}_{A\times A}(\langle 1,2\rangle, \langle 1,3\rangle, \langle 5,6\rangle, \langle 7,8\rangle) = \langle 5,8\rangle,$$

which does not fit the description of $\mathcal{K}$ since, because $\langle 1,2\rangle \neq \langle 1,3\rangle$, the result should be $\langle 7,8\rangle$ rather than $\langle 5,8\rangle$ if $A \times A \in \mathcal{K}$.

5.4 Sets of cardinality $\geq n > 1$

One of the simpler signatures is the following:

sorts: $\{\mathsf{Data}\}$

operators: none

The algebras are, of course, just sets. The algebras form a variety, though not all that interesting a variety. However, if we just add the requirement that they all have a cardinality greater than some number $n > 1$ then they will no longer form a variety because they will not be closed under subalgebras. For that matter, they will also not be closed under products (empty products in particular) nor will they be closed under homomorphic images. So this rather simple example manages to avoid all the Birkhoff conditions. However, if we distinguish between empty and non-empty products we then retain closure under non-empty products. However, if we look at sets of cardinality n for some fixed $n > 1$ then we will also lose closure under non-empty products.

6 Other Specification Techniques

6.1 Conditional axioms (positive Horn formulae)

A conditional axiom (aka positive Horn formula) is of the form

$$e_1 = e'_1 \ \wedge \ \ldots \wedge \ e_n = e'_n \Rightarrow e_{n+1} = e'_{n+1}.$$

Conditional axioms are a strictly more powerful tool than equations for describing categories of Σ-algebras. By taking $n = 0$ we get equations, thus it follows that all varieties can be specified with conditional axioms. On the other hand, it is easy to see that what we called One-to-one algebras can be specified using conditional axioms. The desired conditional axiom is

$$\mathsf{mark}(d) = \mathsf{mark}(d') \Rightarrow d = d'.$$

It follows that categories of algebras specified by conditional axioms are not necessarily closed under homomorphic images.

Clearly though they are closed under products (both empty and non-empty), and under subalgebras. The categories are also abstract and regular. The distinction between regular and closed under homomrophic images is somewhat subtle – recall that

> A category $\mathcal{K}$ is *regular* if for any homomorphism $h : A \rightarrow B$ in $\mathcal{K}$ has a(n "epi-mono") factorization $m \bullet e$ where e is surjective and m is injective with e and m in $\mathcal{K}$. In an abstract category, $\mathcal{K}$ this means that h factors through its image which can be identified with A/Q_h, i.e., we get $h = h^* \bullet h_{Q_h}$ where
>
> $$\begin{array}{c} h_{Q_h} : A \rightarrow A/Q \\ a \mapsto [a]_Q, \end{array}$$
>
> and
>
> $$\begin{array}{c} h^* : A/Q \rightarrow B \\ [a] \mapsto h(a). \end{array}$$

while, on the other hand,

> A category $\mathcal{K}$ is *closed under homomorphic images* if for any homomorphism $h : A \rightarrow B$ with $A \in \mathcal{K}$ but $B \in \mathbf{Alg}_\Sigma$, we have the image of h and the co-restriction of h to that image, are in $\mathcal{K}$.

We leave it to the reader to convince himself that conditional axioms present categories of algebras that are regular even though they need not be closed under homomorphic images.

6.2 General Horn Formulae

A general Horn formula is either a positive Horn formula or of the form

$$\neg(e_1 = e'_1) \ \vee \ \ldots \ \vee \ \neg(e_{n+1} = e'_{n+1})$$

The second form can also be written as

$$e_1 = e_1' \ \wedge \ \ldots \ \wedge \ e_n = e_n' \Rightarrow \neg(e_{n+1} = e_{n+1}')$$

This then means that we can write formulae such as

$$\mathsf{true} \neq \mathsf{false}.$$

which will give us categories of algebras in which the Bool carrier can not collapse to a single point.

It follows that using general Horn formulae we no longer have closure under the empty product. However it can be shown that we have closure under non-empty products – and, furthermore, that this is the most general form of universally quantified sentences for which we have closure under non-empty product – set [2]. Note that it follows from this that we can not specify If-then-else algebras even using general Horn formulae.

In this framework we can define sets with $\geq n > 1$ elements providing we are willing to name n designated elements as constants which we then provide with axioms forcing them to be distinct – this is just the generalization of the Bool example to n elements.

We leave it to the reader to consider the general Horn formula

$$x \neq x.$$

7 Gentzen Formulae

A Gentzen formula is of the form

$$e_1 = e_1' \ \wedge \ \ldots \ \wedge \ e_n = e_n' \Rightarrow e_{n+1} = e_{n+1}' \ \vee \ \ldots \ \vee \ e_{n+p} = e_{n+p}'.$$

Note that if:

- $n = 0$ and $p = 1$ it is an equation.
- $p = 1$ it is a positive Horn formula i.e., a conditional axiom.
- $p = 0$ it is a general Horn formula of the second form. The idea is that $p = 0$ means the conclusion if false, thus for the formula to be true we must have the premise be false as well, but

  $$\neg(e_1 = e_1' \ \wedge \ \ldots \ \wedge \ e_n = e_n')$$

  is equivalent to

  $$\neg(e_1 = e_1') \ \vee \ \ldots \ \vee \ \neg(e_n = e_n'),$$

  (by deMorgan's law).

Thus we can do everything we did before and more, including If-then-else algebras using the two axioms:

$$x_1 = x_2 \Rightarrow \mathsf{if}(x_1, x_2, x_3, x_4) = x_3$$

and

$$\Rightarrow x_1 = x_2 \ \lor \ \mathsf{if}(x_1, x_2, x_3, x_4) = x_4$$

which is easily seen to be equivalent to

$$x_1 \neq x_2 \Rightarrow \mathsf{if}(x_1, x_2, x_3, x_4) = x_4,$$

as well as

$$\mathsf{if}(x_1, x_2, x_3, x_4) \neq x_4 \Rightarrow x_1 = x_2.$$

Using the same technique we can now do the complete version of the Nice-Bool algebras by means of the Gentzen axioms:

$$\mathsf{true} \neq \mathsf{false}$$

and

$$x = \mathsf{true} \ \lor \ x = \mathsf{false}$$

Similar axioms will yield the sets with exactly n elements – but we have to introduce constants, k_i, $i = 1, \ldots, n$, to serve as names for each of the n elements.

It follows, from these examples, that we have now lost closure under non-empty products. This means that the proof of the existence of free algebras no longer works. Indeed it is easy to see that we do not have a free algebra, $F(X)$, even for the special case where $X = \emptyset$, that is we do not have initial algebras.

For example, consider the signature where

sorts: $\{\mathsf{Data}\}$

operators:

a: $\lambda \to \mathsf{Data}$
b: $\lambda \to \mathsf{Data}$
c: $\lambda \to \mathsf{Data}$

axioms:

$\mathsf{a} \neq \mathsf{b}$

$\mathsf{c} = \mathsf{a} \lor \mathsf{c} = \mathsf{b}$

There is no initial algebra in the corresponding category, $\mathcal{K}$. To see this note that that $\mathcal{K}$ contains

- An algebra A with $A_{\mathsf{Data}} = \{\mathsf{a}, \mathsf{b}\}$ and $\mathsf{c}_A = \mathsf{a}$
- An algebra B with $B_{\mathsf{Data}} = \{\mathsf{a}, \mathsf{b}\}$ and $\mathsf{c}_B = \mathsf{b}$

But then there is know way we can have an algebra I with homomorphisms $\alpha : I \to A$ and $\beta : I \to B$ for in I, because of the second axioms, we will have either $\mathsf{c}_I = \mathsf{a}_I$, in which case there is not β, or we will have $\mathsf{c}_I = \mathsf{b}_I$, in which case there will be no α.

However, it is easy to see in this example that the algebras in $\mathcal{K}$ can be partitioned into two subcategories, $\mathcal{K}_A$ and $\mathcal{K}_B$ such that A is initial in $\mathcal{K}_A$ and B is initial in $\mathcal{K}_A$. It can be shown that any category of algebras specified by Gentzen formula can be partitioned into some collection of categories each of which has an initial object – see [3].

We do, however, still retain closure under subalgebras since this property holds for all univeral formulae.

References

[1] Garret Birkhoff. On the structure of abstract algebras. *Proc. Cambridge Phil. Soc.*, 31:433–454, 1935.

[2] P. M. Cohn. *Universal Algebra.* D. Reidel Publishing Company, Dordrecht, Holland, 1981.

[3] Michael Löwe and Uwe Wolter. Parametric algebraic specifications with Gentzen formulas – from quasi-freeness to free functor semantics. *MSCS*, 5:69–112, 1995.