Chapter 2

Equational Logic

2.1 Syntax

2.1.1 Terms and Term Algebras

The natural logic of algebra is equational logic, whose propositions are universally quantified identities between terms built up from variables with operation symbols. For example we defined lattices using terms built up from variables with the operation symbols $\vee$ and $\wedge$. Six equations between such terms expressed the associativity, commutativity, and idempotence of each operation independently, and two more expressed the absorption laws characterizing their relationship. These equations involved terms of height up to two (in associativity and absorption). Equations between terms of yet greater height are deducible from these axioms.

We now formalize and study these notions.

A graded set is a set each element of which is assigned a natural number, its grade. A ***signature*** or ***similarity type*** Σ is a graded set of ***operation symbols***. The grade of an operation symbol is called its ***arity***. We write Σ_n for the subset of Σ consisting of the n-ary operation symbols; thus Σ partitions as the disjoint union $\Sigma = \Sigma_0 + \Sigma_1 + \ldots$. Since every operation symbol is by definition of finite arity, such a signature is sometimes called ***finitary***. We shall not treat here the more general notion of infinitary signature.

For monoids the signature is $\{., 1\}$, with respective arities 2 and 0. Groups ***expand*** the signature with a unary operation of inverse. For lattices, $\Sigma = \{\vee, \wedge\}$, both operations being binary. Semilattices ***reduce*** the signature of lattices to either of $\{\vee\}$ or $\{\wedge\}$.

In the following we fix a signature Σ and a set X of variables, assumed disjoint (no variable is also an operation symbol).

A ***term*** is either a variable, or an n-ary operation symbol of Σ paired with a list of n terms, having a height defined as the least natural number strictly greater than the height of any of its constituent terms.

Examples

1. When Σ is empty, $T(X) = X$ and all terms are of zero height.

2. When X and Σ_0 are empty (no variables or constants), there are no terms of any height: $T(\emptyset) = \emptyset$.

3. For monoids, $\Sigma = \{., 1\}$ consists of a binary operation symbol . and a constant or zeroary operation symbol 1. When X is empty $T(\emptyset)$ contains one term of height 0, namely 1, one term of height 1, namely 1.1, three terms of height 2, namely (1.1).1, 1.(1.1), and (1.1).(1.1), and so on. When X contains $n - 1$ variables

these five forms multiply to respectively n, n^2, n^3, n^3 and n^4 terms.

Writing T_h for the set of terms of height at most h and T for the set of all terms, we may restate this inductive definition set-theoretically as follows.

$$\begin{aligned} T_0 &= X \cup \Sigma_0 \\ T_{h+1} &= \bigcup_{i>0} (\Sigma_i \times T_h^i) \cup T_h \\ T &= \bigcup_{h \geq 0} T_h \end{aligned}$$

When context does not determine X we expand T to $T(X)$, or to $T_\Sigma(X)$ when Σ is not evident either.

An algebra $\mathcal{A} = (A, f_1, f_2, \ldots)$ of type Σ, or Σ-algebra, is a set A together with an interpretation of the operation symbols of Σ as operations on A of the corresponding arity. Algebras of the same type are called ***similar***.

$T_\Sigma(X)$ is made a Σ-algebra $\mathcal{T}_\Sigma(X) = (T_\Sigma(X), \sigma)$ by interpreting each n-ary operation symbol F_i in Σ as the n-ary operation on terms mapping each n-tuple of terms $(t_1, \ldots, t_n)$ to the term $F_i(t_1, \ldots, t_n)$. Here σ denotes a function from Σ to operations on $T_\Sigma(X)$ of the appropriate arity. An algebra of terms constructible in this way is called a *term algebra*.

2.1.2 Free Algebras

In this section we fix a class $\mathcal{C}$ of Σ-algebras. $\mathcal{C}$ may be empty, or consist of the whole class, or anything in between.

A homomorphism of Σ-algebras $(X, f_1, f_2, \ldots), (Y, g_1, g_2, \ldots)$ is a function $h : X \to Y$ such that for all operation symbols F_i of Σ of arity n and for all n-tuples $(x_1, \ldots, x_n)$ over X, $h(f_i(x_1, \ldots, x_n)) = g_i(h(x_1), \ldots, h(x_n))$.

A Σ-algebra $\mathcal{F}$ in the class $\mathcal{C}$ is ***free on*** (or ***freely generated by***) $X \subseteq F$, when for every Σ-algebra $\mathcal{A}$ in $\mathcal{C}$ and function $f : X \to A$, f extends in exactly one way to a homomorphism $h : \mathcal{F} \to \mathcal{A}$. A free algebra on the empty set is called an ***initial algebra*** of $\mathcal{C}$.

Examples

1. When $\Sigma = \emptyset$, the free algebra on X is $(X, \emptyset)$. Every algebra of the empty signature is therefore free, freely generated by itself. In particular the initial $\emptyset$-algebra is the empty set.

2. In the class of monoids, the trivial monoid is clearly initial. The monoid $\mathbb{N} = (N, +, 1)$ of natural numbers under addition is a free monoid generated by the number 1. This monoid is isomorphic to the monoid of one-letter strings under concatenation. The monoid Δ^* of finite strings over an alphabet Δ is a free monoid generated by Δ.

3. In the class of commutative monoids, the commutative monoid $\mathbb{N}^n$ of n-tuples of natural numbers under pointwise addition is a free algebra generated by the n unit vectors $(1, 0, \ldots, 0)$, $(0, 1, \ldots, 0), \ldots, (0, 0, \ldots, 1)$; in particular the upper right quadrant of the lattice points (points with integer coordinates) of the plane constitute the free commutative monoid on two generators.

4. Replacing "monoid" by "group" in example 2, it suffices to replace the natural numbers $\mathbb{N}$ by the integers $\mathbb{Z}$. The appropriate generalization of the monoid Δ^* is to the group $\Delta^\pm$, which can be understood as an infinite tree (connected acyclic undirected graph) for which the set of edges incident on each vertex is $2 \times \Delta$ (so four edges when Δ has two letters). Any vertex of this tree can serve equally well as the interpretation of the unit 1.

5. When X is finite, the semilattice $(2^X - \emptyset, \bigcup)$ is a free semilattice on the set X. When $\emptyset$ is included it becomes the free bounded semilattice (i.e. one having a unit 0). For infinite X we take just the finite nonempty subsets of X, omitting "nonempty" if the unit 0 is in the signature. However $(2^X, \bigcup)$ is a free *complete* semilattice on X.

6. The free distributive lattice on $n = 0, 1, 2$ generators has respectively 0, 4, 18, and 166 elements. These embed as sublattices of free Boolean algebras on the same number of generators, having respectively 2, 16, 256, and 65536 elements. The formula for the latter is well-known to be 2^{2^n}; no closed form is known however for the former.

7. All finite-dimensional vector spaces over the reals are free, freely generated by any choice of basis. This is in fact true for the finite-dimensional vector spaces over any fixed field.

A free algebra in the class $\mathcal{C}$ is naturally referred to as a "free $\mathcal{C}$-algebra." When $\mathcal{C}$ is the class of monoids, groups, or Boolean algebras, their free algebras are called free monoids, free groups, or free Boolean algebras respectively.

Note that $\mathcal{C}$ appears twice in the definition of "free $\mathcal{C}$-algebra:" the free algebra must belong to $\mathcal{C}$, and the algebras $\mathcal{A}$ with respect to which freedom is "measured" must also belong to $\mathcal{C}$.

Theorem 1 *$T_\Sigma(X)$ is a free Σ-algebra freely generated by X.*

Proof: Given an algebra $\mathcal{A} = (A, \Sigma)$ and a function $f : X \to A$, define $h : T(X) \to A$ inductively as follows. Extend h to T_0 by mapping the constant symbols to their interpretations in $\mathcal{A}$. Now assuming h is defined on T_k, namely the set of terms of height at most k, extend it to T_{k+1} by defining h on the terms of height $k + 1$. Since these terms are of the form $F_i(t_1, \ldots, t_n)$ where the t_j's are all in T_k, we may take h of such a term to be $f_i(h(t_1), \ldots, h(t_n))$.

By proceeding systematically by height we ensure that h is defined exactly once on each term. Hence no conflicts arise, so the construction is well-defined. Furthermore every term is catered for, leaving no further freedom for choice of h. ■

Theorem 2 *Free algebras freely generated by isomorphic sets are isomorphic.*

Proof: Each direction of the isomorphism between the sets of generators extends uniquely to a homomorphism between the algebras in that direction. These two homomorphisms compose yielding a homomorphism on one of the two algebras (depending on the order of composition), call it A. This composite extends the identity operation on the generators of A. But the identity homomorphism extends the identity on generators, whence this composite is the identity homomorphism. This is true for either order of composition and so the two homomorphisms between the two algebras constitutes an isomorphism. ■

It follows that in the many places above where we have said "a free algebra on," we could have said "the free algebra on," where uniqueness is understood in the usual sense of algebra as uniqueness up to isomorphism. Furthermore the isomorphism itself is uniquely determined: "the" connotes "unique up to a unique isomorphism."

The following technique permits any algebra to be constructed from a free algebra, provided there are sufficiently large free algebras (the case in most natural situations admitting any free algebras).

Theorem 3 *In any given class $\mathcal{C}$ containing free algebras with arbitrarily large sets of generators, every $\mathcal{C}$-algebra arises as a homomorphic image of a free $\mathcal{C}$-algebra.*

Proof: Given algebra $\mathcal{A}$, let $\mathcal{F}$ be a free algebra generated by a set large enough to have a surjection to the underlying set of A. This surjection extends to a homorphism from $\mathcal{F}$ to $\mathcal{A}$, itself necessarily surjective, which makes $\mathcal{A}$ a homomorphic image of $\mathcal{F}$. ■

Homomorphic images and quotients of an algebra are the same thing up to isomorphism. Hence taking all isomorphic copies of quotients of the free $\mathcal{C}$-algebras, when there are enough of them, yields all of $\mathcal{C}$ (and possibly much more).

It is natural to conjecture that the free algebras of class $\mathcal{C}$ are those satisfying just the equations holding of the class as a whole and no others. The two smallest free monoids in example 2 refute this: the trivial monoid satisfies all equations in the language of monoids, and $\mathbb{N}$ satisfies commutativity.

This sort of counterexample can be accounted for by observing that this condition at least holds for the terms expressible with the available variables. However there is another class of counterexamples to the conjecture, those obtainable by adjoining new elements. For example the free group on n generators satisfies the same monoid equations (those not mentioning inverse) as the free monoid on n generators, yet is not a free monoid.

2.1.3 Equational Logic

Again we assume a fixed signature Σ, in turn fixing the class of Σ-algebras from which all algebras in the sequel are drawn.

An ***equation*** is a pair of terms $s{=}t$ from a common term algebra $T(X)$. (In this section an equal sign between terms means not that the two terms must be the same term but rather denotes their pairing up to form an equation. As a reminder of this the customary space between the terms and the equal sign is omitted.)

Algebra $\mathcal{A}$ ***satisfies*** equation $s{=}t$, written $\mathcal{A} \models s{=}t$, just when for every homomorphism $h : T(X) \to \mathcal{A}$, $h(s) = h(t)$, that is, just when the pair (s, t) is in the kernel of every homomorphism from $T(X)$ to $\mathcal{A}$. That is, no matter what values we assign to the variables of $T(X)$ (defining the homomorphism h), s and t evaluate identically.

We say that a class $\mathcal{C}$ of algebras satisfies an equation just when every algebra in $\mathcal{C}$ satisfies the equation.

The set of all equations on a term algebra $T(X)$ satisfied by some algebra or class is the ***equational theory on*** X ***of*** that algebra or class, denoted $\Theta_X(\mathcal{C})$.

Example. The equational theory of the class containing the one algebra $(\mathbb{Z}, +, \times)$ includes $(x + y, y + x)$, $(x + (y + z), (x + y) + z)$, and $(x(y + z), xy + xz)$.

A ***congruence*** on an algebra $(X, f_1, f_2, \ldots)$ is an equivalence relation $\equiv$ on that algebra such that for every operation f_i of the algebra of arity $n = a_i$ and for every pair $(x_1, \ldots, x_n)$, $(y_1, \ldots, y_n)$ of n-tuples in X^n, if $x_j \equiv y_j$ for $1 \leq j \leq n$ then $f_i(x_1, \ldots, x_n) \equiv f_i(y_1, \ldots, y_n)$.

$\Theta_X(\mathcal{C})$ is a congruence (exercise).

There is one further useful property of equational theories: they are substitutive. In order to define this term we need the notion of a substitution.

A ***substitution*** is a homomorphism on a term algebra.

Informally, a substitution replaces the variables of each term by terms. All occurrences of a variable are replaced by the same term throughout. A variable may be replaced by itself, so a substitution need only change a single variable, or even none at all (the identity substitution).

A binary relation R on a term algebra $T(X)$ is ***substitutive*** when for any substitution h on $T(X)$, if sRt then $h(s)Rh(t)$.

Lemma 4 *Equational theories are substitutive.*

Proof: Let $\mathcal{C}$ be a class having equational theory $\Theta_X(\mathcal{C})$ on $T(X)$ and let $h : T(X) \to T(X)$ be an endomorphism of $T(X)$. If for some pair $s{=}t$ of terms $h(s){=}h(t)$ is not in $\Theta_X(\mathcal{C})$, there must exist A in $\mathcal{C}$ and a homomorphism $g : T(X) \to A$ distinguishing $h(s)$ and $h(t)$. But then the homomorphism $gh : T(X) \to A$ distinguishes s and t, whence $s{=}t$ is not in $\Theta_X(\mathcal{C})$ either. ∎

The ***substitution closure*** of R, written $S(R)$, is the set $\{h(s), h(t) \mid s{=}t \in R \text{ and } h \text{ is a substitution}\}$. $S(R)$ can be seen to be the least substitutive relation containing R.

2.1.4 Exercises

1. When Σ and X are finite, $T_\Sigma(X)$ contains only finitely many terms of a given height.

2. Give a closed-form formula for the number of variable-free monoid terms (i.e. binary trees) of height h. (This is not the Catalan numbers, which go by leaf count, not height.)

3. (a) Prove or refute in each case: the subclass of a class $\mathcal{C}$ consisting of its free algebras is closed under (i) quotients; (ii) subalgebras; (iii) direct products. (b) If one of these three closure properties holds of $\mathcal{C}$, determine whether it implies either of the other two.

4. Given a set X of variables for use in terms, show that the equational theory of a class of Σ-algebras is a congruence on $T_\Sigma(X)$.

5. Show that the intersection of any set of substitutive relations on a term algebra is a substitutive relation on that algebra.

6. Show that the congruence closure of a substitutive relation is substitutive.

2.2 Equational Consequences and Homologues

As we have seen before, in any given state s of affairs a fact p either does or does not hold in that state. This two-valued view of truth determines a binary relation of satisfaction between states and facts. In the previous section, for any given signature Σ we defined the satisfaction relation between the class of all Σ-algebras and the class of all equations between Σ-terms over any given set X of variables.

Now every binary relation between two sets, and more generally classes, determines a concrete Galois connection consisting of polarities between the subclasses of those classes. In the case of satisfaction the polarities arise as the operations $\mathcal{M}$, *models of*, and Θ, *theory of*.

As with any Galois connection, these polarities in turn determine closure operations $\Theta\mathcal{M}$ and $\mathcal{M}\Theta$, understood as respectively *deductive closure* and *homologue closure*. Given any class Γ of equations, $\Theta\mathcal{M}(\Gamma)$ is the class of *equational consequences* of Γ, that is, consequences with respect to equational satisfaction. Similarly given any class $\mathcal{C}$ of algebras, $\mathcal{M}\Theta(\mathcal{C})$ is the class of *equational homologues* of $\mathcal{C}$, those algebras satisfying the equations holding of all the algebras in $\mathcal{C}$.

The following sections characterize the deductive and homologue closure operations of equational logic as respectively substitutive congruence closure, or closure under the standard rules of equational reasoning, and HSP closure, or closure under homomorphic images, subalgebras, and direct products.

2.2.1 Characterization of Deductive Closure

The definition of deductive closure as $\Theta\mathcal{M}$ is a semantic one, inasmuch as it passes through the semantic notion of "all models of Γ." The equivalent characterization we give here is independent of the notions of satisfaction and model, and as such is purely syntactic.

Again we fix a signature Σ, allowing us to abbreviate "Σ-algebra" to just "algebra" and similarly for other signature-dependent notions.

The exercises of the previous section established among other things that for any class $\mathcal{C}$ of algebras, $\Theta(\mathcal{C})$ is a substitutive congruence. In particular this is true of the class $\mathcal{M}(\Gamma)$ of models of a set Γ of equations.

Lemma 5 *Let $\cong$ be any substitutive congruence containing Γ. Then the quotient $T/\cong$ is a model of Γ.*

Proof: Let $h : T \to T/\cong$ be an arbitrary homomorphism from T to $T/\cong$. Take $g : T \to T$ to be the substitution mapping each generator x of T to some element of the equivalence class $h(x)$, and take $q : T \to T/\cong$ to be the evident quotient, namely of T by $\cong$. By construction qg agrees with h on X, whence it must equal h since T is a term algebra. Now if $s{=}t$ is in Γ, it is in $\cong$, whence so is $(g(s), g(t))$ since $\cong$ is substitutive. Hence $q(g(s)) = q(g(t))$ by the definition of q, i.e. $h(s) = h(t)$. Hence $T/\cong$ satisfies every equation in Γ, i.e. it is a model of Γ. ∎

Theorem 6 *For any set Γ of equations, $\Theta(\mathcal{M}(\Gamma))$ is the least substitutive congruence containing Γ.*

Proof: Let $s{=}t$ be in $\Theta\mathcal{M}(\Gamma)$ and let $\cong$ be any substitutive congruence containing Γ. Since $T/\cong$ is a model of Γ, the quotient $q : T \to T/\cong$, being a homomorphism to a model of Γ, must satisfy $q(s) = q(t)$, i.e. s and t are in the same equivalence class of $\cong$, i.e. $s \cong t$. ∎

2.2.2 Equational Reasoning

We now consider the constructive implications of the preceding theorem.

We can form the substitutive congruence closure of Γ in stages by adding equations as described in the following rules.

R1. Infer $x = x$ for any term x.
R2. From $x = y$ infer $y = x$.
R3. From $x = y$ and $y = z$ infer $x = z$.
R4. From $x_1 = y_1, \ldots, x_n = y_n$ infer $f(x_1, \ldots, x_n) = f(y_1, \ldots, y_n)$
for any n-ary operation f of Σ.
R5. From $x = y$ infer $h(x) = h(y)$ where h is a substitution.

The equational consequences of Γ according to this axiom system are those equations deducible from Γ using finitely many applications of rules R1-R5.

Rules R1–R3 give the conditions for the set of all such consequences of Γ to be an equivalence relation. Rule R4 strengthens this to a congruence. Rule R5 further strengthens it to a substitutive congruence.

Theorem 7 *This system of equational reasoning is sound and complete.*

Proof: Any substitutive congruence containing Γ must include all equations deducible from Γ. Hence the equational consequences of Γ obtained via this system must be the least substitutive congruence containing Γ, which we have previously shown to be $\Theta\mathcal{M}(\Gamma)$. ∎

A previous exercise showed that the congruence closure of a substitutive relation is substitutive. It follows that we can get all the substitutions out of the way in the beginning. This convenience may be incorporated into the above axiom system by modifying R5 as follows.

R5'. From axiom $x = y$ infer $h(x) = h(y)$ where h is a substitution.

2.2.3 Characterization of Homologue Closure

A ***variety*** is the class of models of some equational theory. The reader should verify that every variety is closed under homomorphisms, subalgebras, and direct products. Here we prove the converse: if a class C is so closed then it is a variety, i.e. there exists an equational theory Θ such that C is a variety the class of models of Θ. We shall prove this result only for the case where all operations of the signature are finitary, though with care the result can be extended to signatures with infinitary operations, and even to signatures having a proper class of operations whose arities are not bounded by any cardinal.

We prove this in two stages. First we consider the following weaker closure property.

A ***pseudovariety*** is a class closed under isomorphisms, subalgebras, and direct products.

Our strategy will be to show that pseudovarieties contain their free algebras, and then that every algebra of a class C is representable as a quotient of a free algebra of C.

In the following we fix a set X of variables, determining a set $T = T(X)$ of terms.

Lemma 8 *Let $\mathcal{C}$ be a pseudovariety with equational theory $\Theta = \Theta_X(\mathcal{C})$. Then T/Θ belongs to C.*

Proof: Let $E = \overline{\Theta}$ be the set of equations falsifiable by $\mathcal{C}$, i.e. not in Θ. For each $e \in E$ let $A_e \in \mathcal{C}$ and $h_e : T \to A_e$ witness the absence of $e = (s,t)$ from the theory, namely via $h_e(s) \neq h_e(t)$. Define $h : T \to \prod_e A_e$ such that $h(s) = \langle h_e(s) \rangle$. The image $h(T)$ is a subalgebra of $\prod_e A_e$, whence $h(T)$ is in $\mathcal{C}$ by the closure properties of $\mathcal{C}$. But by construction the kernel of h is Θ, whence T/Θ is isomorphic to $h(T)$ and so is in $\mathcal{C}$. ∎

Theorem 9 *(Birkhoff). If $\mathcal{C}$ is closed under homomorphisms, subalgebras, and direct products, then there exists an equational theory Θ on countably many variables such that $\mathcal{C}$ is the class of models of Θ.*

Proof: Take the theory to be $\Theta_N(\mathcal{C})$ on a countable set N of variables. Certainly every member of $\mathcal{C}$ is a model of $\Theta_N(\mathcal{C})$, so it remains to show the converse. Let $\mathcal{A} \models \Theta_N(\mathcal{C})$. Form $T(A)/\Theta_A(\mathcal{C})$, the free algebra in $\mathcal{C}$ on the set A of elements of $\mathcal{A}$. The identity function on A is a function from the set A of generators of $T(A)/\Theta_A(\mathcal{C})$ to the underlying set of the algebra $\mathcal{A}$, and therefore extends to a homomorphism from $T(A)/\Theta_A(\mathcal{C})$ to $\mathcal{A}$. But this homomorphism is onto by construction, making $\mathcal{A}$ a homomorphic image of $T(A)/\Theta_A(\mathcal{C})$. Since $\mathcal{C}$ is closed under homomorphisms it follows that $\mathcal{A}$ is in $\mathcal{C}$. ∎

2.2.4 Exercises

1. Show that $\Theta_X(\mathcal{C})$ is a congruence.

2. Show that if $\equiv$ is a congruence on the algebra $\mathcal{A} = (X, f_1, f_2, \ldots)$ then the quotient $\mathcal{A}/\equiv$ is an algebra.

3. Let $\mathcal{M}_X(\Gamma)$ denote the class of all models of a set Γ of equations between terms of $T(X)$, and let X and Y be two sets of distinct cardinalities. Show that $\mathcal{M}_X\Theta_X(\mathcal{C}) = \mathcal{M}_Y\Theta_Y(\mathcal{C})$ holds for all signatures Σ and classes $\mathcal{C}$ of Σ-algebras if and only if both X and Y are infinite.