Characterizing the Equational Theory

Most of the early work on Kleene algebra was directed toward characterizing the equational theory of the regular sets. These are equations such as $(x+y)^* = x^*(yx^*)^*$ and $x(yx)^* = (xy)^*x$ that hold under the standard interpretation of regular expressions as regular sets of strings.

It turns out that this theory is quite robust and can be characterized in many different ways. We have defined several different but related classes of algebras: Kleene algebras, star-continuous Kleene algebras, closed semirings, and Conway's S-algebras, N-algebras, and R-algebras, all defined axiomatically, as well as relational and language-theoretic algebras defined model-theoretically. All these classes of models have the same equational theory over the signature $+$, $\cdot$, *, 0, 1 of Kleene algebra, and it is the same as the equational theory of the regular sets.

Let us say more carefully what we are talking about. Let σ denote the signature $+, \cdot, ^*, 0, 1$ of Kleene algebra. A *σ-algebra* is any structure of signature σ. This just means that there are distinguished binary operations $+$ and $\cdot$, a distinguished unary operation *, and distinguished constants 0 and 1 defined on C. The structure need not satisfy the axioms of Kleene algebra.

The set of regular expressions RExp_Σ over an alphabet Σ can be regarded as a σ-algebra. The elements of RExp_Σ are just the well-formed terms over variables Σ and operators $+$, $\cdot$, *, 0, 1. The distinguished operations are the syntactic ones; for example, $+$ in RExp_Σ is the binary operation that takes regular expressions s and t and produces the regular expression $s+t$.

For any two σ-algebras C and C', a *homomorphism* from C to C' is a map $h : C \to C'$ that commutes with all the distinguished operations and constants of σ; that is, for all $x, y \in C$,

$$\begin{aligned} h(x+y) &= h(x) + h(y) \\ h(xy) &= h(x) \cdot h(y) \\ h(x^*) &= h(x)^* \\ h(0) &= 0 \\ h(1) &= 1. \end{aligned}$$

Here the operators and constants on the left-hand side are interpreted in C and on the right-hand side in C'. A homomorphism h is

- an *epimorphism* if it is onto; that is, if for all $y \in C'$, there is an $x \in C$ such that $h(x) = y$;
- a *monomorphism* if it is one-to-one; that is, if for all $x, y \in C$, if $h(x) = h(y)$ then $x = y$;
- an *isomorphism* if it is both an epimorphism and a monomorphism.

An *interpretation* is just a homomorphism whose domain is RExp_Σ. For example, let Reg_Σ denote the Kleene algebra of regular sets over alphabet Σ. The *canonical interpretation* over Reg_Σ is the unique homomorphism $R_\Sigma : \mathsf{RExp}_\Sigma \to \mathsf{Reg}_\Sigma$ such that $R_\Sigma(a) = \{a\}$, $a \in \Sigma$. We will show that this interpretation alone characterizes the equational theory of Kleene algebras, as well as all the other classes of algebras mentioned above (Theorem 0.1).

In general, for any σ-algebra C and any function $h : \Sigma \to C$ defined on Σ, h extends uniquely to a homomorphism $\widehat{h} : \mathsf{RExp}_\Sigma \to C$; that is, h and $\widehat{h}$ agree on Σ. Because of this property, the structure RExp_Σ is called the *free σ-algebra on generators* Σ. We say *the* free σ-algebra because it is unique up to isomorphism. Intuitively, once we know how to interpret the letters in Σ, that uniquely determines the interpretation of any regular expression over Σ.

Let s, t be regular expressions and let $I : \mathsf{RExp}_\Sigma \to C$ be an interpretation. We say that the equation $s = t$ *holds under interpretation* I if $I(s) = I(t)$. We say that $s = t$ *holds in* C or that C *satisfies* $s = t$ if $s = t$ holds under all interpretations $I : \mathsf{RExp}_\Sigma \to C$. If $\mathcal{A}$ is a class of algebras or a class of interpretations, we say that $s = t$ *holds in* $\mathcal{A}$ if it holds in all members of $\mathcal{A}$. The *equational theory* of $\mathcal{A}$ is the set of equations that hold in $\mathcal{A}$. The equational theory of $\mathcal{A}$ is denoted $\mathcal{E}(\mathcal{A})$.

Theorem 0.1 *The following classes of algebras all have the same equational theory: Kleene algebras, star-continuous Kleene algebras, closed semirings,* S*-algebras,* N*-algebras,* R*-algebras, language models, and relational models. Moreover, an equation $s = t$ over alphabet Σ is a member of this theory iff it holds under the canonical interpretation $R_\Sigma : \mathsf{RExp}_\Sigma \to \mathsf{Reg}_\Sigma$.*

One can see from this theorem that the equational theory of Kleene algebras is quite robust indeed. If the equational theory were all that we were interested in, there would not be much more to say.

Some Constructions

We will not be able to complete the proof of Theorem 0.1 today. Some parts of the theorem follow immediately from inclusion relationships among the classes of interpretations, but others are more difficult.

First we note that if $\mathcal{A}$ and $\mathcal{B}$ are two classes of algebras or classes of interpretations and $\mathcal{A} \subseteq \mathcal{B}$, then $\mathcal{E}(\mathcal{B}) \subseteq \mathcal{E}(\mathcal{A})$, since any equation that holds in all members of $\mathcal{B}$ must perforce hold in all members of $\mathcal{A}$. We have already established the following inclusions among the classes mentioned in Theorem 0.1:

$$\begin{tikzcd}
\text{Kleene algebras} \arrow[d, dash] & \\
\text{star-continuous Kleene algebras} \arrow[d, dash] \arrow[dr, dash] & \\
\text{closed semirings} \arrow[d, dash] & \mathsf{N}\text{-algebras} \arrow[dl, dash] \arrow[d, dash] \\
\mathsf{S}\text{-algebras} & \text{relational models} \arrow[d, dash] \\
& \text{language models} \arrow[d, dash] \\
& \mathsf{Reg}_\Sigma \arrow[d, dash] \\
& \mathsf{Reg}_\Sigma, R_\Sigma
\end{tikzcd}$$

If class $\mathcal{A}$ occurs above $\mathcal{B}$ in this diagram and there is a path from $\mathcal{A}$ down to $\mathcal{B}$, then $\mathcal{E}(\mathcal{A}) \subseteq \mathcal{E}(\mathcal{B})$. Note that for the two lowest entries in this diagram, the upper one Reg_Σ refers to the equations that hold under any interpretation in Reg_Σ, whereas the lower one $\mathsf{Reg}_\Sigma, R_\Sigma$ refers to the equations that hold under the canonical interpretation only.

First we observe that the equational theories of the S-algebras and the N-algebras coincide. Recall that the N-algebras are the subsets of S-algebras closed under the Kleene algebra operations considered as σ-algebras. We have $\mathcal{E}(\mathsf{N}) \subseteq \mathcal{E}(\mathsf{S})$, since every S-algebra is a subalgebra of itself, therefore is an N-algebra. Conversely, since equations are universal sentences, any equation holding in an S-algebra A holds in any subalgebra of A; therefore $\mathcal{E}(\mathsf{S}) \subseteq \mathcal{E}(\mathsf{N})$.

This observation says that the equational theories of the following classes of interpretations are linearly ordered by inclusion as follows: Kleene algebras, star-continuous Kleene algebras, closed semirings, S-algebras, N-algebras, relational models, language models, Reg_Σ, and $\mathsf{Reg}_\Sigma, R_\Sigma$.

We might also add R-algebras to this list. Recall that R-algebras are those algebras that satisfy all the same equations as N-algebras, thus $\mathcal{E}(\mathsf{R}) = \mathcal{E}(\mathsf{N})$. It will turn out that all the algebras in the diagram above are R-algebras, since they all share the same equational theory, so the class of R-algebras sits at the very top of the diagram above and at the head of the list in the previous paragraph.

However, the concept of R-algebra is not very interesting or useful for axiomatic purposes. Conway [1, p. 102] gives a four-element R-algebra R_4 that is not a star-continuous Kleene algebra. The elements of R_4 are $\{0,1,2,3\}$, and the operations are given by the following tables:

$+$	0	1	2	3
0	0	1	2	3
1	1	1	2	3
2	2	2	2	3
3	3	3	3	3

$\cdot$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	2	3
3	0	3	3	3

$*$	
0	1
1	1
2	3
3	3

To show that R_4 is an R-algebra, it suffices to construct an epimorphism $h : \mathsf{Reg}_\Sigma \to R_4$, since any equation that holds in a σ-algebra also holds in all its homomorphic images. Take $h(0) \stackrel{\text{def}}{=} 0$, $h(1) \stackrel{\text{def}}{=} 1$, and for any other set A,

$$h(A) \stackrel{\text{def}}{=} \begin{cases} 2, & \text{if } A \text{ is finite,} \\ 3, & \text{otherwise.} \end{cases}$$

One can verify easily that this is an epimorphism, therefore R_4 is an R-algebra. It is not a star-continuous Kleene algebra, since $2^n = 2$ for all n, but $2^* = 3$. It is also easily shown that all finite Kleene algebras are star-continuous, therefore R_4 is not a Kleene algebra either.

The family Reg_Σ of regular events over an alphabet Σ gives an example of a star-continuous Kleene algebra that is not a closed semiring. If A is nonregular, the countable set $\{\{x\} \mid x \in A\}$ has no supremum. However, the power set of Σ^* does form a closed semiring.

To construct a closed semiring that is not an S-algebra, we might take the countable and co-countable subsets of ω_1 (the first uncoutable ordinal) with operations of set union for $\sum$, set intersection for $\cdot$, $\varnothing$ for 0, ω_1 for 1, and $A^* = \omega_1$.

To complete the picture, we should construct a Kleene algebra that is not star-continuous. Let ω^2 denote the set of ordered pairs of natural numbers and let $\perp$ and $\top$ be new elements. Order these elements such that $\perp$ is the minimum element, $\top$ is the maximum element, and ω^2 is ordered lexicographically in between. Define $+$ to give the supremum in this order. Define $\cdot$ according to the following table:

$$\begin{aligned} x \cdot \perp &= \perp \cdot x = \perp \\ x \cdot \top &= \top \cdot x = \top, \quad x \neq \perp \\ (a,b) \cdot (c,d) &= (a+c, b+d). \end{aligned}$$

Then $\perp$ is the additive identity and $(0,0)$ is the multiplicative identity. Finally, define

$$a^* = \begin{cases} (0,0), & \text{if } a = \perp \text{ or } a = (0,0) \\ \top, & \text{otherwise.} \end{cases}$$

It is easily checked that this is a Kleene algebra. We verify the axiom

$$ax \leq x \;\; \to \;\; a^* x \leq x$$

explicitly. Assuming $ax \leq x$, we wish to show $a^*x \leq x$. If $a = \perp$ or $a = (0,0)$, then $a^* = (0,0)$ and we are done, since $(0,0)$ is the multiplicative identity. If $x = \perp$ or $x = \top$, we are done. Otherwise, $a > (0,0)$ and $x = (u,v)$, in which case $ax > x$, contradicting the assumption.

This Kleene algebra is not star-continuous, since $(0,1)^* = \top$, but

$$\sum_n (0,1)^n \quad = \quad \sum_n (0,n) \quad = \quad (1,0).$$

References

[1] John Horton Conway. *Regular Algebra and Finite Machines.* Chapman and Hall, London, 1971.