8.5.6 Axioms for string concatenation

In this section we will axiomatize a very simple structure, the structure of string concatenation. A string concatenation system consists of a set A of strings of symbols from some alphabet together with the operation of *concatenation*, which is an operation that applies to two strings and consists simply of writing the second down after the first so as to combine them into a single longer string. In order for the system to be well-defined, the set A of strings must be closed under the concatenation operation; that is, the result of concatenating any two strings in A must itself be in A.

There are two formally different kinds of string concatenation systems, differing in whether they include an empty string among the strings of the system or not We can show how that difference corresponds to a difference of one axiom in otherwise identical axiom systems.

For concatenation systems without an empty string, we can axiomatize them as shown below; structures with a binary operation satisfying these axioms are called *semigroups*

DEFINITION 8.17 *A system consisting of a set A and a binary operation* $\frown$ *on A is* a semigroup *iff:*

1. *A is closed under* $\frown$: $\forall x \forall y((x \in A \,\&\, y \in A) \rightarrow x^\frown y \in A)$

2. *The operation* $\frown$ *is associative:* $\forall x \forall y \forall z((x^\frown y)^\frown z = x^\frown(y^\frown z))$

■

To write these axioms in pure predicate logic form, we would need to eliminate the operator notation "$x^\frown y$". (Similar conversions must be made in going from the function-oriented programming language LISP to the predicate-logic-based language PROLOG.) We can do that by using the notation $Cxyz$ with the intended interpretation "$x^\frown y = z$". That would also force us to stipulate more carefully the existence and uniqueness requirements implicit in the operator notation. The revised first axiom would read as follows:

DEFINITION 8.18 ...

1. *A is closed under C:* $\forall x \forall y \exists z(Cxyz \,\&\, \forall w(Cxyw \rightarrow w = z))$

2. ...

■

Axiom 2 of our earlier definition would also have to be revised, of course, but it merely becomes more complicated and harder to read, so we refrain from carrying out the revision.

An example of a concatenation system of this kind, i.e. a model of the above axioms where the set A is indeed a set of strings and the operation $\frown$ is interpreted as concatenation, is the set of all strings of a's, b's, and c's whose total length is even: $A = \{aa, ab, ac, ba, bb, \ldots, abaa, abab, abac, \ldots, cbccab, \ldots\}$. The set A is closed under concatenation and the concatenation operation is associative.

The set A' which is just like A above except that all the strings in A' have odd length, together with the operation of concatenation, would *not* form a model of the axioms, because it does not satisfy Axiom 1. (Why not?)

Turning now to systems that include the empty string, the first question is what that means. The empty string, like the number zero or the empty set, has more formal than intuitive motivation. It has length zero; it is a substring of every string; and it has the property that when concatenated with any string it yields that string itself. This last is its defining property in the axiomatic characterization of concatenation systems with the empty string: letting e designate the empty string, $x^\frown e = x$, and $e^\frown x = x$, for any string x. The empty string therefore satisfies the definition of being an *identity element* with respect to concatenation, just as 0 is an identity element for addition, 1 is for multiplication, and the empty set is for set union

A concatenation system with empty string therefore satisfies both of the earlier axioms plus an axiom specifying the existence of an identity element; structures that satisfy these axioms are called *monoids*. A monoid is therefore characterizable in general as a semigroup with an identity element

DEFINITION 8.19 *A system consisting of a set A and a binary operation $^\frown$ on A is* a monoid *iff:*

1. *A is closed under $^\frown$:* $\forall x \forall y((x \in A \,\&\, y \in A) \rightarrow x^\frown y \in A)$
2. *The operation $^\frown$ is associative:* $\forall x \forall y \forall z((x^\frown y)^\frown z = x^\frown(y^\frown z))$
3. *A contains an identity element e:* $\exists e \forall x(x^\frown e = e^\frown x = x)$

■

Both monoids and semigroups are examples of kinds of *algebras*. We will return to them in Chapter 10 in the context of group theory and other related algebras. Some parts of the study of algebras relate closely to the study of model theory, since algebras are usually characterizable with a small set of simple axioms whose models can be shown to share rich and significant structural properties. (Among the algebras to be studied in Chapters 9-12, lattices, Boolean algebras, and Heyting algebras have played a particularly important role in model theoretic investigations.)