Term Models & Equational Completeness

We showed in Notes 1 that the equational proof system based on the Ring axioms was *complete*. That is, provability ($\vdash$) and validity ($\models$) coincide for arithmetic equations (aeq's) over the reals or integers.

This connection between provability and validity actually works out just as well, and in a simple way, not only for the Ring axioms, but for the theory of *any* set of equational axioms, as we now show.

We begin by generalizing the definition of arithmetic expressions (ae's) to expressions involving operations on other algebraic structures.

1 First-order Terms

A *signature*, Σ, specifies the names of operations and the number of arguments (*arity*) of each operation. For example, the signature of arithmetic expressions is the set of names $\{\texttt{0}, \texttt{1}, +, \cdot, -\}$ with $+$ and $\cdot$ each of arity two, and $-$ of arity one. The constants $\texttt{0}$ and $\texttt{1}$ by convention are considered to be operations of arity zero.

As a running example, we will consider a signature, Σ_0, with three names: F of arity two, G of arity one, and a constant, c.

The *First-order Terms over* Σ are defined in essentially the same way as arithmetic expressions. We'll omit the adjective "first-order" in the rest of these notes.

Definition 1.1. The set, $\mathcal{T}_\Sigma$, of *Terms* over Σ are defined inductively as follows:

- Any variable, x, is a term over Σ.[1]
- Any constant, $c \in \Sigma$, is a term over Σ.
- If $f \in \Sigma$ has arity $n > 0$, and $M_1, \ldots, M_n \in \mathcal{T}_\Sigma$, then the *application* of f to $M_1, \ldots, M_n$ is in $\mathcal{T}_\Sigma$. We use the usual notation "$f(M_1, \ldots, M_n)$" for this application term.

For example,

$$c, x, F(c, x), G(G(F(y, G(c))))$$

are examples of terms over Σ_0.

[1] We needn't specify exactly what variables are. All that matters is that variables are distinct from other kinds of terms and from operation names in the signature.

2 Substitution

A *substitution* over signature, Σ, is a mapping, σ, from the set of variables to the expressions in $\mathcal{T}_\Sigma$. The notation

$$[x_1, \dots, x_n := M_1, \dots, M_n]$$

describes the substitution that maps variables $x_1, \dots, x_n$ respectively to $M_1, \dots, M_n$, and maps all other variables to themselves.

Definition 2.1. Every substitution, σ, defines a mapping, $[\sigma]$, from $\mathcal{T}_\Sigma$ to $\mathcal{T}_\Sigma$ defined inductively as follows:

$$\begin{aligned} x[\sigma] &::= \sigma(x) && \text{for each variable, } x, \\ c[\sigma] &::= c && \text{for each constant, } c, \\ f(M_1, \dots, M_n)[\sigma] &::= f(M_1[\sigma], \dots, M_n[\sigma]) && \text{for each } f \in \Sigma \text{ of arity } n > 0. \end{aligned}$$

For example,

$$F(G(x), y)[x, y := F(c, y), G(x)] = F(G(F(c, y)), G(x)).$$

3 Models

A *model* assigns meaning to terms by specifying a domain of values and the meaning of the operations in a signature. Models are also called "first-order structures" or "algebras."

Definition 3.1. A *model*, $\mathcal{M}$, for signature, Σ, consists a nonempty set, $\mathcal{D}_\mathcal{M}$, called the *domain*[2] of $\mathcal{M}$, and a mapping that assigns an n-ary operation on the domain to each symbol of arity n in Σ. That is, letting $[\![f]\!]_0$ be the meaning of $f \in \Sigma$, we have for each f of arity $n > 0$,

$$[\![f]\!]_0 : (\mathcal{D}_\mathcal{M})^n \to \mathcal{D}_\mathcal{M},$$

and for each $c \in \Sigma$ of arity 0,

$$[\![c]\!]_0 \in \mathcal{D}_\mathcal{M}.$$

For example, a model for Σ_0 might have domain equal to the set of binary strings, with F meaning the concatenation operation, G meaning reversal, and c meaning the string 001.

Definition 3.2. An $\mathcal{M}$-*valuation*, V, is a mapping from variables into the domain, $\mathcal{D}_\mathcal{M}$.

Once we have a model and valuation, we can evaluate any term to arrive at its value in the domain. The meaning, $[\![M]\!]_\mathcal{M}$, of the term itself is defined to be the function from valuations to the term's value under a valuation. We'll usually omit the subscript when it's clear which model, $\mathcal{M}$, is being referenced.

Definition 3.3. The meaning, $[\![M]\!]$, of term, M, in model, $\mathcal{M}$, is defined by structural induction on the definition of M:

$$\begin{aligned} [\![x]\!]V &::= V(x) && \text{for each variable, } x,. \\ [\![c]\!]V &::= [\![c]\!]_0 && \text{for each constant, } c \in \Sigma, \\ [\![f(M_1, \dots, M_n)]\!]V &::= [\![f]\!]_0([\![M_1]\!]V, \dots, [\![M_n]\!]V) && \text{for each } f \in \Sigma \text{ of arity } n > 0. \end{aligned}$$

[2]The domain is also called the *carrier* of $\mathcal{M}$

There is a fundamental relationship between substitution and meaning given by:

Lemma 3.4 (General Substitution). *Let σ be a substitution and V a valuation. Define V_σ to be the valuation in which*

$$V_\sigma(x) ::= [\![\sigma(x)]\!]V$$

for all variables, x. Then for every term, M,

$$[\![M[\sigma]]\!]V = [\![M]\!]V_\sigma,$$

The proof of Lemma 3.4 follows by structural induction on M as in the Substitution Lemma in Notes 4.

Problem 1. Prove the General Substitution Lemma 3.4.

Definition 3.5. An *equation* consists of two terms called its *lefthand side* and its *righthand side*. The usual notation "$M = N$" is used to denote the equation with lefthand side M and righthand side N. The equation is *valid* in $\mathcal{M}$, written,

$$\mathcal{M} \models M = N,$$

iff $[\![M]\!]_{\mathcal{M}} = [\![N]\!]_{\mathcal{M}}$. If $\mathcal{E}$ is a set of equations, we say that $\mathcal{E}$ is valid in $\mathcal{M}$, written,

$$\mathcal{M} \models \mathcal{E},$$

when $\mathcal{M} \models M = N$ for each equation $(M = N) \in \mathcal{E}$. Finally, we say that $\mathcal{E}$ *logically (or semantically) implies* another set, $\mathcal{E}'$, of equations, written,

$$\mathcal{E} \models \mathcal{E}',$$

when $\mathcal{E}'$ is valid for any model in which $\mathcal{E}$ is valid. That is, for every model, $\mathcal{M}$,

$$\mathcal{M} \models \mathcal{E} \quad \text{implies} \quad \mathcal{M} \models \mathcal{E}'.$$

4 A Proof System for Equations

There are some standard rules for proving equations over a given signature from any set, $\mathcal{E}$, of equations. The equations in $\mathcal{E}$ are called the *axioms*. We write $\mathcal{E} \vdash M = N$ to indicate that equation $(M = N)$ is provable from the axioms. The proof rules are given in Table 1.

We naturally expect of any proof system that what it proves is valid. This property is called *soundness*.

Theorem 4.1 (Soundness). *If $\mathcal{E} \vdash M = N$, then $\mathcal{E} \models M = N$.*

Proof. The Theorem follows by induction on the (tree) structure of the formal proof that $M = N$. The only case that is not completely routine is when $M = N$ is a substitution instance of an axiom. That is, M is $M'[\sigma]$ and N is $N'[\sigma]$ for some axiom $(M' = N') \in \mathcal{E}$ and substitution, σ. But this case

Table 1: Standard Equational Inference Rules.

	$\Longrightarrow$	$M[\sigma]$	$=$	$N[\sigma]$ for $(M = N) \in \mathcal{E}$	(axiom substitution)
	$\Longrightarrow$	M	$=$	M	(reflexivity)
$M = N$	$\Longrightarrow$	N	$=$	M	(symmetry)
$L = M, M = N$	$\Longrightarrow$	L	$=$	N	(transitivity)
$M_1 = N_1, \ldots, M_n = N_n$	$\Longrightarrow$	$f(M_1, \ldots, M_n)$	$=$	$f(N_1, \ldots, N_n)$ when $\text{arity}(f) = n > 0$	(congruence)

follows directly from the Substitution Lemma 3.4. Namely, if $\mathcal{M} \models \mathcal{E}$, then since $(M' = N') \in \mathcal{E}$, we have in particular that $\mathcal{M} \models M' = N'$, that is,

$$[\![M']\!]_{\mathcal{M}} = [\![N']\!]_{\mathcal{M}}. \tag{1}$$

But then,

$$\begin{aligned}
[\![M]\!]V &= [\![M'[\sigma]]\!]V \\
&= [\![M']\!]V_\sigma && \text{(by Subst. Lemma 3.4)} \\
&= [\![N']\!]V_\sigma && \text{(by (1))} \\
&= [\![N'[\sigma]]\!]V && \text{(by Subst. Lemma 3.4)} \\
&= [\![N]\!]V,
\end{aligned}$$

which shows that $[\![M]\!] = [\![N]\!]$, that is, $\mathcal{M} \models M = N$.

So we have shown that if $\mathcal{M} \models \mathcal{E}$, then $\mathcal{M} \models M = N$. That is, $\mathcal{E} \models M = N$, as required.

□

5 Derived Rules

The proof system of Table 1 only allows substitution into the given axioms, $\mathcal{E}$. It would be reasonable have an inference rule allowing substitution into *any* provable equation. But there is no need to add this as an inference rule because the current set of rules already justifies such substitutions.

Lemma 5.1 (Provable Substitution). *For any substitution, σ, if $\mathcal{E} \vdash M = N$, then $\mathcal{E} \vdash M[\sigma] = N[\sigma]$.*

The proof of Lemma 5.1 follows by induction on the (tree) structure of the proof that $M = N$, as in the proof of Lemma 3.3 in Notes 4.

Problem 2. Prove the Provable Substitution Lemma.

We can also derive a more general congruence rule. Namely, let σ_1 and σ_2 be substitutions and define

$$\mathcal{E} \vdash \sigma_1 = \sigma_2$$

to mean that $\mathcal{E} \vdash \sigma_1(x) = \sigma_2(x)$ for all variables, x.

Lemma 5.2 (Provable Congruence). *If* $\mathcal{E} \vdash \sigma_1 = \sigma_2$*, then* $\mathcal{E} \vdash M[\sigma_1] = M[\sigma_2]$.

A proof by structural induction on M is routine.

Problem 3. (a) Suppose that the (congruence) rule was replaced by the (Provable Congruence) rule of Lemma 5.2. Show that (congruence) would then be a derived rule in this new proof system.

(b) Prove the (Provable Congruence) rule.

6 Completeness

We are now ready to prove

Theorem 6.1 (Completeness). $\mathcal{E} \models M = N$ *iff* $\mathcal{E} \vdash M = N$.

The right-to-left implication of Theorem 6.1 is just the Soundness Theorem 4.1.

To prove the left-to-right implication, we must show that if an equation $(M = N)$ is not provable, then it is not true in some model that satisfies all the equations in $\mathcal{E}$. We will actually by construct a single model, $\mathcal{M}_{\mathcal{E}}$, which satisfies $\mathcal{E}$ but in which *every* unprovable equation is actually false.

In particular, we will show that

Proposition 6.2.

$$\mathcal{E} \vdash M = N \quad \textit{iff} \quad \mathcal{M}_{\mathcal{E}} \models M = N.$$

Completeness follows directly from Proposition 6.2: since every equation in $\mathcal{E}$ is provable using the (axiom substitution) rule — because an equation is a substitution instance of itself — the left-to-right direction of Proposition 6.2 implies that

$$\mathcal{M}_{\mathcal{E}} \models \mathcal{E}.$$

And if $(M = N)$ is not provable from $\mathcal{E}$, then the right-to-left direction of Proposition 6.2 implies that $(M = N)$ is not valid in $\mathcal{M}_{\mathcal{E}}$.

So all we need to do is construct the model, $\mathcal{M}_{\mathcal{E}}$, and prove Proposition 6.2.

We begin by observing that the proof rules of (reflexivity), (symmetry) and (transitivity) imply that for any fixed $\mathcal{E}$, provable equality between terms M and N is an equivalence relation. We let $\langle M \rangle_{\mathcal{E}}$ be the equivalence class of M under provable equality, that is,

$$\langle M \rangle_{\mathcal{E}} ::= \{N \mid \mathcal{E} \vdash M = N\}.$$

So we have by definition

$$\mathcal{E} \vdash M = N \qquad \text{iff} \qquad \langle M \rangle_{\mathcal{E}} = \langle N \rangle_{\mathcal{E}} . \tag{2}$$

The domain of $\mathcal{M}_{\mathcal{E}}$ will be defined to be the set of $\langle M \rangle_{\mathcal{E}}$ for $M \in \mathcal{T}_{\Sigma}$. So each element of the domain is a set of terms in $\mathcal{T}_{\Sigma}$. That's why the model we construct is called a *term model*.

The meaning of constants, $c \in \Sigma$, will be

$$[\![c]\!]_0 ::= \langle c \rangle_{\mathcal{E}} .$$

The meaning of operations $f \in \Sigma$ of arity $n > 0$ will be determined by the condition

$$[\![f]\!]_0(\langle M_1 \rangle_{\mathcal{E}}, \ldots \langle M_n \rangle_{\mathcal{E}}) ::= \langle f(M_1, \ldots, M_n \rangle_{\mathcal{E}} .$$

Notice that $[\![f]\!]_0$ applied to the equivalence classes $\langle M_1 \rangle_{\mathcal{E}}, \ldots \langle M_n \rangle_{\mathcal{E}}$ is defined in terms of the designated terms $M_1, \ldots, M_n$ in these classes. To be sure that $[\![f]\!]_0$ is well-defined, we must check that the value of $[\![f]\!]_0$ would not change if we chose other designated terms in these classes. That is, we must verify that

$$\text{if } \langle M_1 \rangle_{\mathcal{E}} = \langle N_1 \rangle_{\mathcal{E}}, \ldots, \langle M_n \rangle_{\mathcal{E}} = \langle N_n \rangle_{\mathcal{E}}, \text{ then } \langle f(M_1, \ldots, M_n) \rangle_{\mathcal{E}} = \langle f(N_1, \ldots, N_n) \rangle_{\mathcal{E}} .$$

But this is an immediate consequence of the (congruence) rule. So $\mathcal{M}_{\mathcal{E}}$ is well-defined.

Now for any substitution, σ, let W_σ be the $\mathcal{M}$-valuation given by

$$W_\sigma(x) ::= \langle \sigma(x) \rangle_{\mathcal{E}}$$

for all variables, x. The following key property of the term model now follows by essentially the same structural induction on M used to prove the General Substitution Lemma 3.4.

Lemma 6.3.

$$[\![M]\!] W_\sigma = \langle M[\sigma] \rangle_{\mathcal{E}} . \tag{3}$$

Problem 4. Prove Lemma 6.3.

Now let ι be the identity substitution $\iota(x) ::= x$ for all variables, x. For any term, M, the substitution instance $M[\iota]$ is simply identical to M, so Lemma 6.3 immediately implies

$$[\![M]\!] W_\iota = \langle M \rangle_{\mathcal{E}} .$$

In particular, if $[\![M]\!] = [\![N]\!]$, then $\langle M \rangle_{\mathcal{E}} = \langle N \rangle_{\mathcal{E}}$, so from equivalence (2), we conclude Proposition 6.2 in the right-to-left direction:

Lemma 6.4. *If* $[\![M]\!]_{\mathcal{M}_{\mathcal{E}}} = [\![N]\!]_{\mathcal{M}_{\mathcal{E}}}$*, then* $\mathcal{E} \vdash M = N$*.*

Finally, for the left-to-right direction of Proposition 6.2, we prove

Lemma 6.5. *If* $\mathcal{E} \vdash M = N$*, then* $[\![M]\!]_{\mathcal{M}_{\mathcal{E}}} = [\![N]\!]_{\mathcal{M}_{\mathcal{E}}}$*.*

Proof. If $\mathcal{E} \vdash M = N$, then by (provable substitution), $\mathcal{E} \vdash M[\sigma] = N[\sigma]$ for any substitution, σ. So

$$\langle M[\sigma]\rangle_{\mathcal{E}} = \langle N[\sigma]\rangle_{\mathcal{E}} \tag{4}$$

by (2). Now let V be any $\mathcal{M}$-valuation, and let σ be any substitution such that $\sigma(x) \in V(x)$ for all variables, x. This ensures that

$$V = W_{\sigma}, \tag{5}$$

and we have

$$\begin{aligned}
[\![M]\!]V &= \langle M[\sigma]\rangle_{\mathcal{E}} && \text{(by (5) \& Lemma 6.3)}\\
&= \langle N[\sigma]\rangle_{\mathcal{E}} && \text{(by (4))}\\
&= [\![N]\!]V && \text{(by (5) \& Lemma 6.3).}
\end{aligned}$$

Since V was an arbitrary valuation, we conclude that $[\![M]\!] = [\![N]\!]$, as required. □