2. Equational Theories

2.1. Syntax

Def. 2.1.1: *Let $\mathcal{S}$ be a non-empty set (of* **sorts***). A* **type** *(over $\mathcal{S}$) is a finite sequence of sorts; so Θ, the* **type language***, is defined as $\Theta = \mathcal{S}^+$. A type $T = S_1 \cdots S_n S \in \Theta$ is usually written as $T = S_1 \times \cdots \times S_n \to S$. We say that $S_1 \cdots S_n$ is the* **definition type** *of T and S is its* **image type***. n is the* **arity** *of T.* □

Example 2.1.2: As a set of sorts we take

$$\mathcal{S}^0 \;=\; \{\ \text{INTEGER},\ \text{BOOLEAN}\ \}\ .$$

Types over $\mathcal{S}^0$ are for example:

$$\begin{aligned} T^0 &= \text{INTEGER}, \\ T^1 &= \text{INTEGER} \times \text{INTEGER} \to \text{INTEGER}, \\ T^2 &= \text{INTEGER} \times \text{INTEGER} \to \text{BOOLEAN}. \end{aligned}$$
□

Def. 2.1.3: *A* **signature** *over (the set of sorts) $\mathcal{S}$ is a pair (Σ, r), where Σ is a set (of* **operators** *or* **function symbols***), and r is a* **typing function** *$r : \Sigma \to \Theta$.*
Often we will assume the typing function r to be implicitly given, and we simply speak of the signature Σ.
If $r(C) = S \in \mathcal{S}$, then C is called a **constant** *(operator).* □

Example 2.1.4: (Example 2.1.2 cont.)
$\Sigma^0 \;=\; \{\mathbf{0}, \mathbf{SUCC}, \mathbf{PLUS}, \mathbf{TRUE}, \mathbf{FALSE}, \mathbf{NE}\ \}$, together with

$$\begin{aligned} &r(\mathbf{0}) \;=\; \mathbf{INTEGER}, \\ &r(\mathbf{SUCC}) \;=\; \mathbf{INTEGER} \to \mathbf{INTEGER}, \\ &r(\mathbf{PLUS}) \;=\; \mathbf{INTEGER} \times \mathbf{INTEGER} \to \mathbf{INTEGER}, \\ &r(\mathbf{TRUE}) \;=\; r(\mathbf{FALSE}) \;=\; \mathbf{BOOLEAN}, \\ &r(\mathbf{NE}) \;=\; \mathbf{INTEGER} \times \mathbf{INTEGRE} \to \mathbf{BOOLEAN}, \end{aligned}$$

is a signature over $\mathcal{S}^0$. □

Def. 2.1.5: *Let Σ be a signature over $\mathcal{S}$. A* **Σ-algebra** *is a pair $(\mathcal{A}, \mathcal{F})$, where $\mathcal{A}$ is a $\mathcal{S}$-indexed family of sets, i.e. $\mathcal{A} = \{A_S | S \in \mathcal{S}\}$, and $\mathcal{F}$ is a*

Σ-indexed family of functions, i.e. $\mathcal{F} = \{\mathcal{F}_F | F \in \Sigma\}$, *s.t.*

$$\text{if } r(F) = S \text{ then } \mathcal{F}_F \in \mathcal{A}_s, \text{ and}$$
$$\text{if } r(F) = S_1 \times \cdots \times S_n \to S \text{ then } \mathcal{F}_F = \mathcal{A}_{S_1} \times \cdots \times \mathcal{A}_{S_n} \to \mathcal{A}_S.$$

$\mathcal{A}_S$ is the **carrier set** *or* **universe** *of sort S. $\mathcal{F}_F$ is the* **operation** *or* **function** *of the algebra associated with the operator or function symbol F.*
Often we denote a Σ-algebra simply by its carrier $\mathcal{A}$. Often we will also simply write F instead of $\mathcal{F}_F$. □

Example 2.1.6: (Example 2.1.4 cont.)
We get a Σ^0-algebra $(\mathcal{A}^0, \mathcal{F}^0)$ by setting

$$\mathcal{A}^0_{\text{INTEGER}} := \mathbb{N}, \qquad \mathcal{A}^0_{\text{BOOLEAN}} := \mathbb{B} = \{\texttt{true}, \texttt{false}\},$$

$\mathcal{F}^0(\mathbf{0})$ is (the integer constant) 0,
$\mathcal{F}^0(\mathbf{SUCC})$ is the successor function in $\mathbb{N}$,
$\mathcal{F}^0(\mathbf{PLUS})$ is the addition function in $\mathbb{N}$,
$\mathcal{F}^0(\mathbf{TRUE})$ is (the boolean constant) `true`,
$\mathcal{F}^0(\mathbf{FALSE})$ is (the boolean constant) `false`,
$\mathcal{F}^0(\mathbf{NE})$ is the function from $\mathbb{N}^2$ to $\mathbb{B}$ which returns `true` if and only if the arguments are different. □

Based on a signature Σ we will now define terms as the basic building block of equations. So we will create a Σ-algebra simply from the syntactic material available in the signature. Furthermore, we need to introduce suitable notation for referring to subterms.

Def. 2.1.7: *Let Σ be a signature over $\mathcal{S}$. We consider the following Σ-algebra (with carrier) $\mathcal{T}(\Sigma)$:*

- *if $C \in \Sigma$ is a constant of sort S, then C is an element of $\mathcal{T}(\Sigma)_S$,*
- *if $F \in \Sigma$ with $r(F) = S_1 \times \cdots \times S_n \to S$ and $t_i \in \mathcal{T}(\Sigma)_{S_i}$ for all $1 \le i \le n$, then $F\, t_1 \,\ldots\, t_n$ is an element of $\mathcal{T}(\Sigma)_S$,*
- *nothing else is in $\mathcal{T}(\Sigma)$.*

For every $F \in \Sigma$ the function $\mathcal{F}_F$ takes $t_1, \ldots, t_n$ and produces $F\, t_1 \,\ldots\, t_n$. For better readability we will often write $F(t_1, \ldots, t_n)$ instead of $F\, t_1 \,\ldots\, t_n$. The algebra $\mathcal{T}(\Sigma)$ is called the **algebra of ground terms** *or the* **initial**

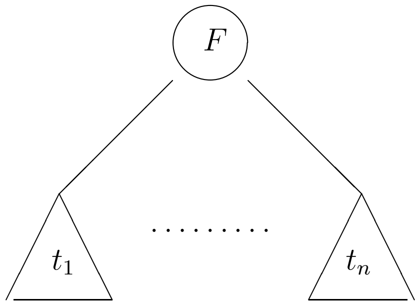

Figure 1: isomorphism term ↔ tree

algebra *over* Σ.
The carrier set $\mathcal{T}(\Sigma)_S$ *is called the* **set of terms** *of sort* S. □

The algebra of ground terms is isomorphic (by a Σ-isomorphism, as will be introduced later) to the algebra of trees over the signature Σ; this isomorphism relates a term

$$F(t_1, \ldots, t_n)$$

with the tree as shown in Figure 1; i.e. the corresponding tree has a root labeled with F and subtrees corresponding to the subterms $t_1, \ldots, t_n$. In computer science we typically speak of "abstract syntax trees", whereas in algebra we often speak of "words".

Next we introduce (general) terms, which are constructed from ground terms and variables.

Def. 2.1.8: *Let* $\mathcal{V}$ *be a* $\mathcal{S}$*-indexed family of sets* $\mathcal{V}_S$*. The elements of* $\mathcal{V}_S$ *are called* **variables** *of sort* S*. We assume* $\mathcal{V}_S \cap \mathcal{V}_{S'} = \emptyset$ *for* $S \neq S'$ *and* $\mathcal{V}_S \cap \Sigma = \emptyset$ *for* $S \in \mathcal{S}$.
By $\Sigma \cup \mathcal{V}$ *we denote the signature which we get by adding to* Σ *every element of* $\mathcal{V}_S$ *as a constant of sort* S.
The resulting algebra $\mathcal{T}(\Sigma \cup \mathcal{V})$ *is called the* **free** Σ**-algebra** *generated by* $\mathcal{V}$*, or the* **term algebra** *over* Σ *and* $\mathcal{V}$.
The carrier set $\mathcal{T}(\Sigma \cup \mathcal{V})_S$ *is called the* **set of terms** *of sort* S. □

If the signature and the set of variables is clear from the context, then we simply write $\mathcal{G}$ for $\mathcal{T}(\Sigma)$ and $\mathcal{T}$ for $\mathcal{T}(\Sigma \cup \mathcal{V})$.

Example 2.1.9: (Example 2.1.6 cont.)
In the initial algebra or algebra of ground terms $\mathcal{T}(\Sigma^0)$ we have, for exam-

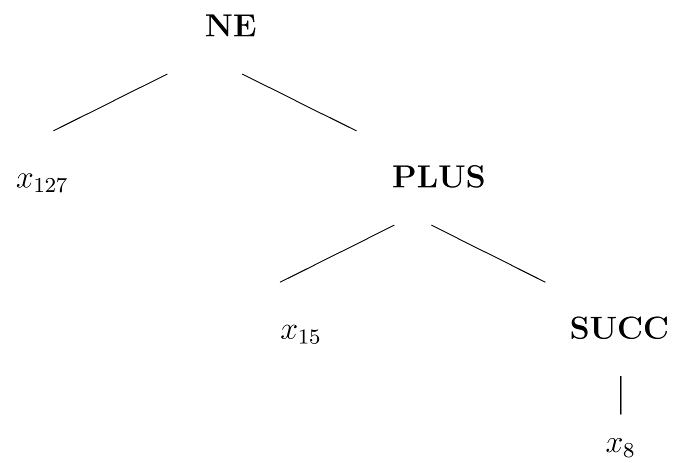

Figure 2: isomorphic tree in Example 2.1.9

ple, the ground terms

$\mathbf{0}$	of sort **INTEGER**
$\mathbf{SUCC}(\mathbf{PLUS}(\mathbf{0}, \mathbf{SUCC}(\mathbf{0})))$	of sort **INTEGER**
TRUE	of sort **BOOLEAN**
$\mathbf{NE}(\mathbf{PLUS}(\mathbf{0}, \mathbf{SUCC}(\mathbf{0})), \mathbf{0})$	of sort **BOOLEAN**

If we take

$$\mathcal{V}^0_{\mathbf{INTEGER}} = \{x_0, x_1, x_2, \ldots\}$$

as variables of sort **INTEGER** and

$$\mathcal{V}^0_{\mathbf{BOOLEAN}} = \{y_0, y_1, y_2, \ldots\}$$

as variables of sort **BOOLEAN** — so $\mathcal{V}^0$ is the family consisting of $\mathcal{V}^0_{\mathbf{INTEGER}}$ and $\mathcal{V}^0_{\mathbf{BOOLEAN}}$ — then the term algebra $\mathcal{T}(\Sigma^0 \cup \mathcal{V}^0)$ contains, for instance, the terms

$\mathbf{0}, x_5$	of sort **INTEGER**
$\mathbf{SUCC}(\mathbf{PLUS}(\mathbf{x_7}, \mathbf{SUCC}(\mathbf{0})))$	of sort **INTEGER**
$\mathbf{NE}(x_{127}, \mathbf{PLUS}(x_{15}, \mathbf{SUCC}(x_8)))$	of sort **BOOLEAN** .

The last term corresponds to the tree in Figure 2. □

Def. 2.1.10: *Let t be a term in a term algebra $\mathcal{T}$. The set of* **occurrences** *or* **positions** *in t is the following subset of $\mathbb{N}^*$, the set of finite sequences of natural numbers:*

$$\mathrm{occ}(t) := \begin{cases} \{\Lambda\}, & \text{if } t \text{ is a variable or a constant,} \\ \{\Lambda\} \cup \{i \cdot p \,|\, 1 \le i \le n, p \in \mathrm{occ}(t)\}, & \text{if } t = F(t_1, \ldots, t_n) \ . \end{cases}$$

By Λ we denote the empty sequence, and "$\cdot$" denotes the concatenation of sequences (so $\Lambda \cdot p = p = p \cdot \Lambda$).
Now suppose that $p_1, p_2, q \in \mathrm{occ}(t)$. By $\leq$ we denote the **prefix ordering** *on $\mathbb{N}^*$; i.e.*

$$p_1 \ \leq \ p_2 \qquad \text{iff} \qquad p_2 = p_1 \cdot p' \text{ for some } p' \in \mathbb{N}^* \ .$$

If $p_1 \leq p_2$, then by p_2/p_1 we mean the sequence p', for which $p_2 = p_1 \cdot p'$; p' is the **quotient** *of p_2 by p_1.*
p_1 and p_2 are **disjoint** *or* **perpendicular**, *$p_1 \perp p_2$, iff $p_1 \not\leq p_2$ and $p_2 \not\leq p_1$.*
□

Def. 2.1.11: *Let t, s be terms in $\mathcal{T}$, and $p \in \mathrm{occ}(t)$. The* **subterm** *of t* **at** *p is*

$$t_p \ := \ \begin{cases} t & \text{if } p = \Lambda \ , \\ (t_i)_q & \text{if } p = i \cdot q \text{ for } i \in \mathbb{N}, q \in \mathbb{N}^* \text{ and } t = F(t_1, \ldots, t_n) \ . \end{cases}$$

The set $\mathcal{V}(t)$ of **variables occurring in** *t is*

$$\mathcal{V}(t) \ := \ \{x \in \mathcal{V} \mid x = t_{/p} \text{ for a } p \in \mathrm{occ}(t)\} \ .$$

The result of the **replacement of the subterm of** *t* **at** *p* **by** *s is defined as*

$$t[p \leftarrow s] \ := \ \begin{cases} s & \text{if } p = \Lambda \ , \\ F(t_1, \ldots, t_{i-1}, t_i[q \leftarrow s], t_{i+1}, \ldots, t_n) & \\ \quad \text{if } p = i \cdot q \text{ for some } i \in \mathbb{N}, q \in \mathbb{N}^*, \text{ and } t = F(t_1, \ldots, t_n). \end{cases}$$
□

Example 2.1.12: Set $\Sigma^0, \mathcal{V}^0$ be as in Example 2.1.9.
$t = \mathbf{NE}(x_{127}, \mathbf{PLUS}(x_{15}, \mathbf{SUCC}(x_8)))$ is a term in $\mathcal{T}(\Sigma^0 \cup \mathcal{V}^0)$.
$\mathrm{occ}(t) = \{\Lambda, 1, 2, 2 \cdot 1, 2 \cdot 2, 2 \cdot 2 \cdot 1\}$.
We have $2 \cdot 2 \leq 2 \cdot 2 \cdot 1$, $2 \cdot 2 \cdot 1/2 \cdot 2 = 1$, and $1 \perp 2 \cdot 2$.
Also $t_{2 \cdot 2} = \mathbf{SUCC}(x_8)$ and $\mathcal{V}(t) = \{x_8, x_{15}, x_{127}\}$.
If we let $s := \mathbf{PLUS}(0, x_1)$, then $t[2 \cdot 2 \leftarrow s] = \mathbf{NE}(x_{127}, \mathbf{PLUS}(x_{15}, \mathbf{PLUS}(0, x_1)))$. □

We mention some simple relation between these notions. Proofs are mere technicalities.

Lemma 2.1.13: *Let s, t be terms in $\mathcal{T}$, and $p, q \in \mathbb{N}^*$.*

(i) *If* $p \cdot q \in \mathrm{occ}(s)$*, then* $q \in \mathrm{occ}(s_p)$ *and* $s_{/p\cdot q} = (s_p)_q$.

(ii) *If* $p \in \mathrm{occ}(s)$ *and* $q \in \mathrm{occ}(s_p)$*, then* $p \cdot q \in \mathrm{occ}(s)$. □

Lemma 2.1.14: (properties of replacement) *Let* $s, t, u \in \mathcal{T}$*,* $p, p_1, p_2 \in \mathrm{occ}(s)$*,* $q \in \mathrm{occ}(t)$*.*

(i) (embedding) $s[p \leftarrow t]_{p\cdot q} = t_q$.

(associativity) $s[p \leftarrow t][p \cdot q \leftarrow u] = s[p \leftarrow t[q \leftarrow u]]$.

(ii) *Let* $p_1 \perp p_2$.

(persistence) $s[p_1 \leftarrow t]_{p_2} = s_{p_2}$.

(commutativity) $s[p_1 \leftarrow t][p_2 \leftarrow u] = s[p_2 \leftarrow u][p_1 \leftarrow t]$.

(iii) *Let* $p_2 \leq p_1$.

(distributivity) $s[p_1 \leftarrow t]_{p_2} = (s_{p_2})[p_1/p_2 \leftarrow t]$.

(dominance) $s[p_1 \leftarrow t][p_2 \leftarrow u] = s[p_2 \leftarrow u]$.

Proof of (embedding): by induction on the length of p.
If $p = \Lambda$, then $s[\Lambda \leftarrow t]_{\Lambda\cdot q} = t_{/q}$.
Now assume (induction hypothesis) that the assertion holds for some $\tilde{p}$, and then consider $p = i \cdot \tilde{p}$, for some $i \in \mathbb{N}$. Let $s = F(s_1, \ldots, s_i, \ldots, s_n)$. We get

$$
\begin{array}{lll}
s[p \leftarrow t]_{p\cdot q} & = & s[i \cdot \tilde{p} \leftarrow t]_{i\cdot\tilde{p}\cdot q} \\
 & = & F(s_1, \ldots, s_i[\tilde{p} \leftarrow t], \ldots, s_n)_{i\cdot\tilde{p}\cdot q} \\
 & = & s_i[\tilde{p} \leftarrow t]_{\tilde{p}\cdot q} \\
 & =_{\text{(ind.hyp.)}} & t_q \ .
\end{array}
$$

This completes the proof. □

Def. 2.1.15: *Let* $\mathcal{A}, \mathcal{B}$ *be* Σ*-algebras. Let* $h : \mathcal{A} \to \mathcal{B}$ *be an* $\mathcal{S}$*-indexed family of mappings* $h_S : \mathcal{A}_S \to \mathcal{B}_S$*.*
Then we say that h *is a* Σ-**morphism** *from* $\mathcal{A}$ *to* $\mathcal{B}$ *iff for all* $F \in \Sigma$ *with* $r(F) = S_1 \times \cdots \times S_n \to S$ *we have*

$$
h_S(F_{\mathcal{A}}(a_1, \ldots, a_n)) \ = \ F_{\mathcal{B}}(h_{S_1}(a_1), \ldots, h_{S_n}(a_n)) \ . \qquad \square
$$

Def. 2.1.16: *Let $\mathcal{A}$ be a Σ-algebra. A mapping $\nu : \mathcal{V} \to \mathcal{A}$ is called an* **evaluation function** *over $\mathcal{A}$. (Actually ν is an $\mathcal{S}$-indexed family of mapping $\nu_S : \mathcal{V}_S \to \mathcal{A}_S$, for $S \in \mathcal{S}$.)* □

Theorem 2.1.17 (freeness of the term algebra)**:** *Let $\mathcal{A}$ be a Σ-algebra. Every evaluation function $\nu : \mathcal{V} \to \mathcal{A}$ can be extended uniquely to a Σ-morphism from $\mathcal{T}(\Sigma \cup \mathcal{V})$ to $\mathcal{A}$.* □

Def. 2.1.18: *A* **substitution** *is a Σ-endomorphism σ on $\mathcal{T}$, such that $\sigma(x) = x$ for almost all (i.e. all but finitely many) variables.*
$\mathrm{D}(\sigma) := \{x \in \mathcal{V} \mid \sigma(x) \neq x\}$ *is called the* **domain** *of σ.*
σ is a **ground substitution** *iff $\mathcal{V}(\sigma(x)) = \emptyset$ for all $x \in \mathrm{D}(\sigma)$.* □

A substitution σ is characterized by the corresponding set

$$\{\sigma(x) \to x \mid x \in \mathrm{D}(\sigma)\}.$$

Example 2.1.19: (a) Let $\mathcal{A}^0$ be the Σ^0-algebra of Example 2.1.6.
We get another Σ^0-algebra $\mathcal{A}^1$ by setting
$\mathcal{A}^1_{\text{INTEGER}} := \mathbb{Z}_5$, $\mathcal{A}^1_{\text{BOOLEAN}} := \mathbb{B} = \{\texttt{true}, \texttt{false}\}$,
$\mathcal{F}^1(\mathbf{0})$ is 0,
$\mathcal{F}^1(\mathbf{SUCC})$ is the successor function modulo 5,
$\mathcal{F}^1(\mathbf{PLUS})$ is the addition function modulo 5,
$\mathcal{F}^1(\mathbf{TRUE})$ is (the boolean constant) `true`,
$\mathcal{F}^1(\mathbf{FALSE})$ is (the boolean constant) `false`,
$\mathcal{F}^1(\mathbf{NE})$ is the function which always returns `false`.
The mapping $h : \mathcal{A}^0 \to \mathcal{A}^1$ with

$$h_{\text{INTEGER}}(m) = m \mod 5, \quad \text{and} \quad h_{\text{BOOLEAN}}(b) = \texttt{false}$$

is a Σ^0-morphism from $\mathcal{A}^0$ to $\mathcal{A}^1$.
(b) Let $\mathcal{V}^0$ be as in Example 2.1.9. Then the mapping $\nu : \mathcal{V}^0 \to \mathcal{A}^0$ with

$$\nu_{\text{INTEGER}}(x_i) = i \quad \text{and} \quad \nu_{\text{BOOLEAN}}(y_i) = \begin{cases} \texttt{true} & \text{for } i \text{ even} \\ \texttt{false} & \text{for } i \text{ odd} \end{cases}$$

is an evaluation function over $\mathcal{A}^0$.
If we extend ν to a Σ^0-morphism from $\mathcal{T}(\Sigma^0 \cup \mathcal{V}^0)$ to $\mathcal{A}^0$, then we get, for

instance,

$$\begin{aligned}
&\nu(\mathbf{NE}(x_{127}, \mathbf{PLUS}(x_{15}, \mathbf{SUCC}(x_8))) \ = \\
&\mathbf{NE}_{\mathcal{A}^0}(\nu(x_{127}), \mathbf{PLUS}_{\mathcal{A}^0}(\nu(x_{15}), \mathbf{SUCC}_{\mathcal{A}^0}(\nu(x_8)))) \ = \\
&\mathbf{NE}_{\mathcal{A}^0}(127, \mathbf{PLUS}_{\mathcal{A}^0}(15, 9))) \ = \\
&\mathbf{NE}_{\mathcal{A}^0}(127, 24) \ = \\
&\texttt{true}\ .
\end{aligned}$$

□

Lemma 2.1.20: *Let $s, t \in \mathcal{T}$, σ a substitution, $p \in \mathrm{occ}(s)$. Then we have the following:*

(i) $\mathrm{occ}(\sigma(s)) \ = \ \mathrm{occ}(s) \ \cup \ \bigcup_{\substack{q \in \mathrm{occ}(s) \\ s_q \in \mathcal{V}}} \{q \cdot q' \mid q' \in \mathrm{occ}(\sigma(s_q))\}.$

(ii) $\sigma(s)_p = \sigma(s_p)$.

(iii) *If $s_p \in \mathcal{V}$, then $\sigma(s)_{pq} = \sigma(s_p)_q$ for all $q \in \mathrm{occ}(\sigma(s_p))$.*

(iv) $\sigma(s[p \leftarrow t]) = \sigma(s)[p \leftarrow \sigma(t)]$.

Def. 2.1.21: *Let $s, t \in \mathcal{T}$. Then t is an* **instance** *of s, written as $s \preceq t$, iff $\sigma(s) = t$ for a substitution σ.*
$\preceq$ is a partial ordering on $\mathcal{T}$, the **subsumption ordering**.
If $s \preceq t$ and $s \neq t$ then we write $s \prec t$. □

Lemma 2.1.22: *$\prec$ is a Noetherian relation on $\mathcal{T}$; i.e., there is no finite sequence of term $t_0, t_1, \ldots$ such that $\cdots \prec t_1 \prec t_0$.* □

Now we are prepared to speak about equations and equational theories. For equational theories we introduce a proof calculus, the equational calculus. This is nothing else but a restriction of inference rules in 1st order predicate calculus to the situation of equational axioms with only universal quantification (which is not explicity written) and "=" as the only predicate. This system is also called equational logic.

Definition 2.1.23: *A* Σ*–***equation** *(or* **equation** *for short) is a pair* $s = t$*, such that* $s, t \in \mathcal{T}_S$ *for a sort* $S \in \mathcal{S}$*.*
Let E *be a set of equations and* $s, t \in \mathcal{T}$*. Then* s *and* t *are* **provably equal** *modulo* E*, or* $s = t$ **is provable** *from* E*, written as* $E \vdash s = t$*, iff* $s = t$ *can be derived from* E *in finitely many steps in the following* **equational calculus***:*

(G1) *elements of* E *are axioms:*

$$\frac{}{u_1 = u_2}$$

for all $u_1 = u_2 \in E$

(G2) *reflexivity, symmetry, and transitivity:*

$$\frac{}{u_1 = u_1}, \quad \frac{u_1 = u_2}{u_2 = u_1}, \quad \frac{u_1 = u_2, u_2 = u_3}{u_1 = u_3}$$

for all $u_1, u_2, u_3 \in \mathcal{T}$

(G3) *substitution rule:*

$$\frac{u_1 = u_2}{\sigma(u_1) = \sigma(u_2)}$$

for all $u_1, u_2 \in \mathcal{T}$*,* σ *a substitution*

(G4) *replacement rule:*

$$\frac{u_1 = u_1', \ldots, u_n = u_n'}{F(u_1, \ldots, u_n) = F(u_1', \ldots, u_n')}$$

for all $u_1, \ldots, u_n, u_1', \ldots, u_n' \in \mathcal{T}$*,* $F \in \Sigma$ *with appropriate type*

The **equational theory** $=_E$ *generated by* E *consists of all equations, which can be proven from* E*:*

$$=_E \quad = \quad \{\ s = t \ \mid \ E \vdash s = t\ \}.$$

We also use the notation $s =_E t$ *instead of* $E \vdash s = t$*.*
E *is called a* **basis** *for the equational theory* $=_E$*.* □

Compare [BN98], p.42.

Def. 2.1.24: *Let $\mathcal{A}$ be a Σ-algebra. A sort-preservind equivalence relation $\sim$ on $\mathcal{A}$ is called a* **Σ-congruence** *on $\mathcal{A}$ iff*

$$\big(\forall F \in \Sigma \text{ of definition type } S_1 \dots S_n\big)$$
$$\big(\forall a_1, b_1 \in \mathcal{A}_{S_1}, \dots, a_n, b_n \in \mathcal{A}_{S_n}\big)$$
$$a_1 \sim b_1, \dots, a_n \sim b_n \quad \Longrightarrow \quad F(a_1, \dots, a_n) \sim F(b_1, \dots, b_n)\ . \ \square$$

Theorem 2.1.25: *Let E be a set of equations. Then $=_E$ is the weakest Σ-congruence over the term algebra $\mathcal{T}$, which contains all pairs $\sigma(s) = \sigma(t)$ for $s = t \in E$ and σ a substitution.* $\square$

Proof: Let

$$\mathcal{K} := \bigcap \{\ K \mid \ K \text{ is a } \Sigma-\text{congruence over } \mathcal{T} \text{ containing all } \sigma(s) = \sigma(t) \text{ for } s = t \in E \text{ and } \sigma \text{ a substitution } \}\ .$$

It is clear that $\mathcal{K}$ is a Σ-congruence, and it is the weakest Σ-congruence with these properties.
We show that $=_E = \mathcal{K}$.

Because of (G2) the relation $=_E$ is an equivalence relation.
Now consider $F \in \Sigma$ with definition type $S_1 \dots S_n$, $u_i, u_i' \in \mathcal{T}_{S_i}$ and $u_i =_E u_i'$ for $1 \leq i \leq n$. Because of (G4) we have $F(u_1, \dots, u_n) = F(u_1', \dots, u_n')$. So $=_E$ is a Σ-congruence, and therefore

$$=_E \supseteq \mathcal{K}\ . \tag{1}$$

We show "$=_E \subseteq \mathcal{K}$ by induction on the length l of the proof for $s =_E t$. If $l = 1$, the only step in the proof must be an application of (G1) or of the first part of (G2). In both cases $s = t \in \mathcal{K}$.
In the induction hypothesis we assume that for every equation $s =_E t$ having a proof of length $< l$, the pair (s, t) is in $\mathcal{K}$.
Now let $s =_E t$ have a proof of length l. If the last proof step is one of (G1), (G2), or (G4), then by inspection we see that also (s, t) has to be in $\mathcal{K}$.
Finally we have to show that also the application of rule (G3) does not lead out of $\mathcal{K}$. We consider the modified equational basis

$$E' := \{\, \sigma(s) = \sigma(t) \mid s = t \in E \text{ and } \sigma \text{ a substitution} \,\}\ .$$

Obviously

$$=_{E'} = =_E\ ,$$

so it suffices to show that $=_{E'} \subseteq \mathcal{K}$.
In the same way as above we see that (G1,2,4) do not lead out of $\mathcal{K}$. But in a proof modulo E' there is no need to ever use rule $(G3)$. Consider a shortest proof modulo E', in which we use (G3), say

$$P: \quad \begin{array}{l} s_1 = t_1 \\ \vdots \\ s_j = t_j \\ \vdots \\ s_n = t_n \end{array}$$

The last proof step being an application of (G3), there must be a $j < n$ such that $s_n = \sigma(s_j), t_n = \sigma(t_j)$. Now instead of every axiom e in P we could use the axiom $\sigma(e)$. Thus in proof step j we would get $\sigma(s_j) = \sigma(t_j)$ without ever having used rule (G3). So

$$=_E \, = \, =_{E'} \subseteq \mathcal{K} \tag{2}$$

and this completes the proof. □

2.2. Semantics

Def. 2.2.1: *Let $\mathcal{A}$ be a Σ-algebra and $s = t$ an equation. Then we call $s = t$* **valid** *in $\mathcal{A}$, or $\mathcal{A}$ a* **model** *of $s = t$, and we write $\mathcal{A} \models s = t$ or $s =_{\mathcal{A}} t$, iff $\nu(s) = \nu(t)$ for every evaluation function $\nu : \mathcal{V}(s) \cup \mathcal{V}(t) \to \mathcal{A}$.*
If E is a set of equations, then E is **valid** *in $\mathcal{A}$, or $\mathcal{A}$ is a* **model** *of E, iff $\mathcal{A} \models e$ for every $e \in E$.* □

Example 2.2.2: The equation $x_1 + x_2 = x_2 + x_1$ is valid both in $\mathcal{A}^0$ (see Example 2.1.6) and also in $\mathcal{A}^1$ (see Example 2.1.19).
The equation $x_1 + x_1 + x_1 + x_1 + x_1 = 0$ is not valid in $\mathcal{A}^0$ but it is valid in $\mathcal{A}^1$. □

Def. 2.2.3: (a) *If E is a set of equations, then by $\mathcal{M}(E)$ we denote the class of all models of E. $\mathcal{M}(E)$ is called the* **variety** *of E.*
(b) *If $\mathcal{C}$ is a class of algebras (over the same signature Σ), then the* **validity problem** *for $\mathcal{C}$ asks for a decision of $C \models e$ for arbitrary $C \in \mathcal{C}$ and Σ-equation e. If this is the case, then we say that e is* **valid** *in $\mathcal{C}$ and we write $\mathcal{C} \models e$.*
If $\mathcal{C}$ consists of only one algebra C, then we also speak of the **word problem** *over C.*
(c) *Let E be a set of equations and $s = t$ an equation. Then we say that $s = t$* **follows** *from E, or s and t are* **semantically equal** *modulo E, and we write $E \models s = t$, iff $\mathcal{M}(E) \models s = t$; i.e., $s = t$ is valid in every model of E.* □

Example 2.2.4: We consider the axioms for groups:

$$G: \quad \begin{array}{l} 1 \cdot x = x \ , \\ x^{-1} \cdot x = 1 \ , \\ (x \cdot y) \cdot z = x \cdot (y \cdot z) \ . \end{array}$$

Then we have

$$G \ \models \ x^{-1} \cdot ((y^{-1} \cdot y) \cdot x) = z^{-1} \cdot z \ .$$

We can see this in the following way:
Let $\mathcal{A}$ be an algebra with a constant **one**, a unary operation **inv** and a binary operation **times**.

Assume that $\mathcal{A} \in \mathcal{M}(G)$; so for all $a, b, c \in \mathcal{A}$ we have:

$$\begin{array}{l}\mathbf{times}(\mathbf{one}, a) = a \ , \\ \mathbf{times}(\mathbf{inv}(a), a) = \mathbf{one} \ , \\ \mathbf{times}(\mathbf{times}(a, b), c) = \mathbf{times}(a, \mathbf{times}(b, c)) \ .\end{array}$$

Now for an arbitrary evaluation function ν we have

$$\begin{array}{l}\nu(x^{-1} \cdot ((y^{-1} \cdot y) \cdot x)) \ = \\ \mathbf{times}(\mathbf{inv}(\nu(x)), \mathbf{times}(\mathbf{times}(\mathbf{inv}(\nu(y)), \nu(y)), \nu(x))) \ = \\ \mathbf{times}(\mathbf{inv}(\nu(x)), \mathbf{times}(\mathbf{one}, \nu(x))) \ = \\ \mathbf{one} \ = \\ \mathbf{times}(\mathbf{inv}(\nu(z)), \nu(z)) \ = \\ \nu(z^{-1} \cdot z) \ .\end{array}$$

□

In general, the notions of "provability" and "validity" do not coincide in a (1st order) logical theory. However, in equational logic, which is a very restricted form of 1st order predicate logic, they do coincide. This has been proven by G.Birkoff.

Theorem 2.2.5: (G. Birkhoff, see [B35] [1]) *Let E be the basis of an equational theory. Then*

$$E \models s = t \quad \Longleftrightarrow \quad E \vdash s = t \ .$$

Proof: see J.Avenhaus, "Reduktionssysteme", p. 86. □

[1] [B35] G. Birkhoff, "On the structure of abstract algebras", *Proc. Cambridge Phil. Soc.* 31, pp. 433–454 (1935)