Final Algebra Semantics and Data Type Extensions*

MITCHELL WAND

Computer Science Department, Indiana University, Bloomington, Indiana 47405
Received August 23, 1977; revised February 14, 1979

We consider the problem of data type extensions. Guttag, Horowitz, and Musser have pointed out that in this situation the naive initial algebra approach requires the data type to save too much information. We formulate a category of implementations of such an extension, and we show that such a category has a final object. The resulting semantics is closer to that of Hoare, since it can be argued that an abstract data type in the sense of Hoare is a final object in the category of representations of that type. We consider as an example the specification of integer arrays, and we show that our specification yields arrays as its abstract data type. The connection with initial algebra semantics is discussed.

0. PROLOGUE

In this paper we are concerned with the definition of new data types from old, using the viewpoint of what is called initial algebra semantics [9, 10, 13, 14]. Before discussing the problem in detail, we summarize our interpretation of the initial algebra approach in this prologue.¹

One wishes to specify data types axiomatically, that is, by writing down, in some logical calculus, sentences which describe those properties of the data type on which its user may rely. A program which uses a data type may then be proved correct by deducing its verification conditions from the axioms of the data type. Such a program will then work correctly with any implementation of the data type which satisfies the axioms. Thus the programmer is concerned not with single algebras, but with the class of algebras which are legal representations of the data type; the programs he writes ought to work satisfactorily regardless of which representation is used. Our first thesis, therefore, is that a specification of a data type should present a class of algebras. If one desired merely to construct a single algebra, then numerous mathematical techniques are available; it is the finite presentation of classes of algebras that requires formal methods.

One logical language which seems to be useful for the specification of data types is the language of generators and relations [9, 10, 13, 16]. A presentation via generators and relations defines an equational class of algebras. Since one wishes to discuss connections

* Research reported herein was supported in part by the National Science Foundation under grant number MCS75-06678A01.

¹ The reader should be warned that we diverge in some details from the approach, say, of [10]. Our outlook is much closer to that of [9]; any misinterpretations are solely our responsibility, of course.
between equational classes independent of their presentation, one introduces categories called \textbf{algebraic theories} [20]. An algebraic theory is a representative of its equational class, just as a Zermelo-Fraenkel ordinal is a representative of its order-isomorphism class.\(^2\) An algebraic theory consists of equivalence classes of terms (compositions of generators), where two terms are equivalent iff their equality is deducible from the relations.\(^3\) Of particular importance is the case where there are no relations between the generators; then the theory is called a \textbf{free theory} and the morphisms are just the terms. The \(T\)-algebras (or "implementations", or "models") of an algebraic theory \(T\) are certain functors from \(T\) to the category of sets; this picture is merely a notational variant of the conventional picture of an algebra.

The denotational semantics of a term in a \(T\)-algebra is the mapping it induces on the universe set of the algebra. This mapping is obtained by mapping the term (a morphism of a free theory \(F\)) to its equivalence class (a morphism of \(T\)), and thence, via the \(T\)-algebra (qua functor) to the desired set map (a morphism of the category of sets). We identify \(T\)-algebras with implementations, and since the functor from the free theory to \(T\) is independent of the implementation, we sometimes refer to it as "the semantics". These relationships are shown in Figure 0.1\(^4\).

\[
\begin{array}{c}
F \xrightarrow{\text{semantics}} T \\
\downarrow \text{implementation} \\
\text{Sets}
\end{array}
\]

\textbf{Figure 0.1}

The fragment of algebra semantics we have described is more than an algebraicization of attribute grammars [19] with only synthesized attributes.\(^5\) The difference is that the algebraic framework allows additional problems to be attacked:

(i) \textbf{Equivalence of presentations:} given two sets of generators and relations, do they define the same class of implementations? If the generator sets are quite different, it may be difficult to state a translation theorem; it may be easier to prove the algebraic theories isomorphic. Similar questions arise with respect to simulability [8] or program transformations [3, 23].

(ii) \textbf{Operational semantics:} given some complex term and some set of terms which we regard as known constants, a computation is a deduction (in some appropriate formal system) that the value of the complex term is always equal to the value of a particular known term in any implementation, i.e. that they are mapped by the semantics to the same morphism in \(T\). It can be shown that the problem of whether two terms are equal under \(T\) is equivalent to a word problem in a tree rewriting system [22], and then under

\(^2\) Although, unlike the ordinal, an algebraic theory is \textit{not} a member of the class it represents.

\(^3\) More precisely, the \textbf{Morphisms} of a theory are equivalence classes of (tuples of) terms.

\(^4\) This is, of course, a crude notion of implementation—see (iii) below.

\(^5\) Although synthesized attributes are enough, see, for example [4].
reasonable conditions the tree rewriting system has the Church–Rosser property with various pleasant consequences.\(^6\)

(iii) Classes of implementations: An equational class is usually not quite what one wants for the class of implementations. One may desire additional closure properties (which leads to the consideration of “theories with additional structure” [5, 24]) or more restricted closure properties (a situation to be considered in this paper). Rather than having a single, so-called “abstract” implementation, one always has a class of implementations, and one may pose the question of which of those implementations is “the” desired one. The conventional choice is the initial $T$-algebra, which has two desirable properties. First, its universe contains no values other than those required by the generators. Second, two values have the same semantics in the initial $T$-algebra if and only if they have the same semantics in every $T$-algebra. Thus no information is lost except that which is required by the relations.

Guttag et al. [17] have suggested that in some cases the initial algebra saves too much information. It is the purpose of this paper to suggest a solution to that problem. The outline of the paper is as follows: Section 1 presents an example to illustrate the problem posed by Guttag et al. Section 2 is given over to definitions, most of which are quite standard. In Section 3, we argue that an abstract data type in the sense of [18] ought to be a final object in the category of data type representations. In Section 4, we present our model of data type extensions. In Section 5, we prove the main result: that the category of representations of a data type extension has a final object, which gives the final algebra semantics of the title. It is also shown that the conventional initial algebra semantics is preserved as a special case. We mention briefly an analogy between final algebra semantics and minimal realization in automata.

1. INTRODUCTION

In this section we presume a general familiarity with the mathematical structures discussed in the prologue. There are several excellent tutorials on various aspects of this material [11, 12, 17].

Let us consider a theory of integers, $T_{\mathbb{Z}}$. This theory will have one sort, denoted $i$, and generators as follows:

- for each nonnegative integer $k$, a symbol $n_k : A \to i$
- undefined: $A \to i$
- plus: $i + i \to i$

\(^6\) See [17] for a well-illustrated discussion. Although tree rewriting systems have been the object of some study [22], their exact connection with algebraic theories has not to our knowledge been adequately explored in print.
subject to the identities

\[
\text{plus}[n_k, n_p] = n_{k+p} \text{ for each } k, p \in \omega^7
\]
\[
\text{plus}[x, \text{undefined}] = \text{undefined}
\]
\[
\text{plus}[\text{undefined}, x] = \text{undefined}
\]

It is easy to show, using canonical term algebras [10], that the initial algebra of \( T_{\omega^+} \) consists of \( \{n_k \mid k \in \omega\} \cup \{\text{undefined}\} \) under the usual addition. Now let us consider the theory of a data type which consists of integer-valued arrays indexed by integers (where "integers" is the type defined by \( T_{\omega^+} \)). To do this, we add a new sort (for arrays), called \( a \), and new generators:

- \textbf{empty}: \( A \rightarrow a \) (the empty array)
- \textbf{alt}: \( ai \rightarrow a \) (\text{alt}[A, j, x] = "A after } A[j]: = x")
- \textbf{val}: \( ai \rightarrow i \) (\text{val}[A, j] = A[j])

The intended semantics of the generators (sketched above) may be captured by adding the following identities:

\[
\text{val[empty, } x\text{]} = \text{undefined}
\]
\[
\text{val[alt}[x, n_p, z], n_p] = z \text{ for } p \in \omega
\]
\[
\text{val[alt}[x, n_k, z], n_p] = \text{val}[x, n_p] \text{ for } k \neq p, k, p \in \omega
\]
\[
\text{val}[x, \text{undefined}] = \text{undefined}
\]

We call this theory \( T_{\text{ARR}} \).

It is clear that \( \text{val} \), applied to an array and an integer, always reduces to an integer by applications of the identities, i.e., this set of identities is sufficiently-complete in the sense of Guttag [16]. The initial algebra of \( T_{\text{ARR}} \), however, does not consist of the arrays we hoped to define. The initial algebra consists of two sorts. The sort \( S_i \) corresponding to \( i \) consists of \( \{n_k \mid k \in \omega\} \cup \{\text{undefined}\} \) as before, but the set \( S_a \) corresponding to \( a \) is defined inductively as

(i) \text{empty} \in S_a

(ii) if \( x \in S_a \), and \( m, m' \in S_i \), then \( \text{alt}[x, m, m'] \) is in the set \( S_a \)

(iii) nothing else.

For example

\[
\text{alt[alt[empty, } n_1, n_3\text{], } n_2\text{, } n_2\text{]}
\]

and

\[
\text{alt[alt[empty, } n_2, n_3\text{], } n_1\text{, } n_1\text{]}
\]

This represents a countable set of axioms.
are distinct elements of the initial algebra of $T_{ARR}$. Here the initial algebra saves too much information: it saves not only the values in the array but also the order of all changes in the array.

One could destroy this unneeded information by adding the identities

\[
\begin{align*}
\text{alt}[\text{alt}[x, n_k, y], n_k, z] &= \text{alt}[x, n_k, z] & k \in \omega \\
\text{alt}[\text{alt}[x, n_k, y], n_p, z] &= \text{alt}[\text{alt}[x, n_p, z], n_k, y] & k, p \in \omega, k \neq p \\
\text{alt}[x, \text{undefined}, y] &= x
\end{align*}
\]

It is straightforward to see that this suppresses duplicate subscript entries and causes subscript errors on updates to be ignored. Unfortunately, adding the second axiom scheme causes the underlying operational semantics to lose the Church–Rosser property \cite{17, 22}. This is an unpleasant consequence; Guttag et. al., suggest the use of "equality interpretations" to allow information to be lost in a controlled manner.

It is the purpose of this paper to suggest another solution. We observe that the difficulty arises when we are dealing with data type extensions. We have "enough information" in our implementation of the extension so long as no values of the base type (e.g., integers) are merged. We wish to lose as much information as possible; therefore we are led to final algebras in the category of implementations which have "enough information". The main theorem of this paper shows that such final algebras exist.

2. Preliminaries

If $C$ is a category, $C(a, b)$ denotes the set of arrows or morphisms from object $a$ to object $b$. If $f \in C(a, b)$ and $g \in C(b, c)$, their composition, a member of $C(a, c)$, is denoted $g \cdot f$. We write $gf$ when no confusion results. If $f \in C(a, b)$ then $\text{dom}(f) = a$ and $\text{cod}(f) = b$.

$\text{Sets}$ will denote the category whose objects are sets and whose morphisms are the usual set-theoretic functions. Right-to-left composition (usually of functors or of functions in $\text{Sets}$) is written using "o": $g \circ f(x) = g(f(x))$.

If $C$ is a category, an object $a$ of $C$ is initial iff for any object $b$ of $C$, there is exactly one morphism in $C$ from $a$ to $b$. The object $a$ is final iff for any object $b$ there is exactly one morphism in $C$ from $b$ to $a$. All initial objects in a category are always isomorphic; similarly for final objects. In $\text{Sets}$, $\varnothing$ is initial (consider the function whose graph is empty), and any singleton set is final.

Let $S$ be a set whose elements are called sorts. An $S$-sorted operator alphabet $\Omega$ is a map $\Omega: K \to S^* \times S$ for some set $K$. If $s \in K$, and $\Omega s = (w, a)$, we say $w$ is the domain of $s$ and $a$ is the codomain of $s$. If $S$ has only one element, and $w = a^a$ (where $S = \{a\}$), we say $s$ is n-ary; $\Omega$ is then a ranked alphabet. When no ambiguity results, we will write $\Omega$ for $K$ and write "$s \in \Omega"$. We write $\Omega(w, a)$ for $\{s \in K | \Omega s = (w, a)\}$.

An $S$-sorted algebraic theory (or just theory) is a category $T$ whose objects are the elements of $S^*$ and in which multiplication in $S^*$ coincides with the categorical product.

\[S^*\text{ denotes the free monoid generated by } S. \]
If $T$ is a theory, and $f_i \in T(u, w_i)$ (for $i = 1, \ldots, n$), then the product morphism in $\cdot T(u, w_1 \cdot \cdot \cdot w_n)$ is denoted $[f_1, \ldots, f_n]$. We write $e_i$ for the projection morphisms. A \textit{theory-functor} is a product-preserving functor between theories. If $\Omega$ is an $S$-sorted operator alphabet, we may construct the free theory $F_\Omega$ by the usual methods \cite{12}; if $s \in \Omega$, then $s \in F_\Omega(\text{dom}(s), \text{cod}(s))$.

If $T$ is an $S$-sorted theory, so is $T^2$, where $T^2(u, v) = \{(f, g) \mid f, g \in T(u, v)\}$ with composition given by $(f, g)(f', g') = (ff', gg')$. An \textit{equation} on $T$ is an element of $T^2(w, a)$ for some $a \in S$. A \textit{congruence} on $T$ is a subtheory $R$ of $T^2$ such that for each $u, v \in S^*$, $R(u, v)$ is an equivalence relation on $T(u, v)$. If $R$ is a congruence on $T$, then we can form the quotient theory $T/R$ via $T/R(u, v) = T(u, v)/R(u, v)$. $T/R$ is also an $S$-sorted theory; it is the coequalizer of the evident diagram $R \to T^2 \Rightarrow T$.

If $A$ is a set of equations on $T$, we can construct the smallest congruence on $T$ containing $A$ as the set of theorems of a formal system $E_A$. The formal objects of $E_A$ are the morphisms of $T^2$. We write $(f, f') : u \to v$ for $(f, f') \in T^2(u, v)$, and $\vdash (f, f') : u \to v$ if $(f, f')$ is provable in $E_A$. The axioms and rules of $E_A$ are as follows:

\textbf{Axioms:} \quad If $(f, f') : w \to a \in A$, then $\vdash (f, f') : w \to a$ \textit{EA}

\quad For any $f \in T(u, v)$, $\vdash (f, f) : u \to v$ \textit{ER}

\textbf{Rules:}

\begin{align*}
(f, g) : u \to v & \text{ ES} & (f, g) : u \to v & (g, h) : u \to v & \text{ ET} \\
(g, f) : u \to v & \text{ EC} & (g, f) : w \to y & (f, f') : v \to w & (h, h) : u \to v \\
(g \cdot f \cdot h, g \cdot f' \cdot h) : u \to y & \text{ EP} & (f_1, f'_1) : m \to a_1, \ldots, (f_n, f'_n) : m \to a_n \\
([f_1, \ldots, f_n], [f'_1, \ldots, f'_n]) : m \to a_1 \cdots a_n & \text{ EP}
\end{align*}

Let $E_A(u, v) = \{(f, f') \mid \vdash (f, f') : u \to v\}$. Axiom scheme $E_A$ ensures that every equation in $A$ is in $E_A$; rules $ER, ES, \text{ and } ET$ ensure that each $E_A(u, v)$ is an equivalence relation; rule $EC$ closes $E_A$ to a subcategory of $T^2$, and rule $EP$ closes $E_A$ under the product operation of $T^2$. Hence $E_A$, with composition inherited from $T^2$, is the smallest congruence on $T$ containing $A$.

A theory may be presented by $(\Omega, A)$ where $\Omega$ is an operator alphabet (the \textit{generators}) and $A$ is a set of equations (the \textit{relations}). $(\Omega, A)$ presents the theory $T$ where $T(u, v) = F_\Omega(u, v)/E_A(u, v)$. The functor $F : F_\Omega \to T$ sending each morphism to its equivalence class is a full theory functor.

If $T$ is an $S$-sorted theory, a \textit{T-algebra} is a product-preserving functor $A : T \to \text{Sets}$. A natural transformation $h : A \to B$ from one $T$-algebra to another is just a homomorphism of algebras (over $\Omega$). The $T$-algebras and natural transformations form a category $T\text{-Alg}$.

If $T$ is an $S$-sorted theory, the $T$-algebra $A$ given by

\begin{align*}
A(w) &= T(A, w) \quad w \in S^* \\
A(f) : T(A, s) \to T(A, v) &\quad g \mapsto fg \quad f \in T(w, v)
\end{align*}
is initial in $T$-$Alg$. This (when decoded) comes out to be the conventional term algebra in the case where $T$ is a free theory; where $T$ is not free, the carriers consist of equivalence classes (under $E_d$) of (tuples of) terms. We refer to this particular initial algebra as the canonical initial algebra. The $T$-algebra $Z$ given by $Z(w) = \{I\}$ is final in $T$-$Alg$. $Z$ is the algebra whose universes consist of singleton sets for each sort (and whose operations are therefore trivial).

3. Data Type Representations

One anomalous property of the initial algebra approach is a seeming incompatibility with other notions of abstract data types e.g. [18]. If $A$ is an initial algebra of $T$, and $B$ is any other $T$-algebra, there is a unique morphism $A \rightarrow B$. In Hoare’s version (and in related work [e.g. 21]), the map runs the other way: one has the “abstraction map” from an arbitrary implementation $B$ to the set of “abstract values”. In this section we will attempt to make some sense of these two views.

We said previously that we identify objects of the category $T$-$Alg$ with implementations of the theory $T$. This identification is, of course, rough at best; for example, it includes the final algebra $Z$ as a legal implementation. Even if we wish to exclude some elements of $T$-$Alg$, the class of legal implementations of $T$ will be some subcategory $K$ of $T$-$Alg$.

Let us imagine, therefore, that we are given a particular subcategory $K$ of $T$-$Alg$ which is known to be the category of legal implementations of $T$; and let $W$ be the “abstract data type”. In the example of Section 1, $W$ would be given by

$$W(i) = \{n_k \mid k \in \omega\} \cup \{\text{undefined}\} \quad \text{(as before)}$$

$$W(a) = \{M \mid M \text{ is a partial function } \omega \rightarrow \omega, \text{ of finite domain}\}$$

$$W(\text{val}) = \lambda(M, j) \begin{cases} \text{undefined} & \text{if } j = \text{undefined} \\ \text{undefined} & \text{else if } M \text{ is undefined at } j \\ n_{M(0)} & \text{else} \end{cases}$$

$$W(\text{alt}) = \lambda(M, j, x) [M - \{(j, y) \in M\} \cup \{(j, x)\}]$$

$K$ is to be the category of representations of $W$.

If we have a reasonable notion of “category of legal implementations”, the following observations should hold:

1. $W$ is an object of $K$ (A data type ought to be a legal implementation of itself)
2. for any object $A$ of $K$, there is a morphism in $K$ from $A$ to $W$ (the “abstraction map”)
3. for any object $A$ of $K$, there is only one morphism in $K$ from $A$ to $W$. (There is only one “reasonable” abstraction map for each data type representation $A$, i.e., each “concrete” value in $A$ may reasonably represent only one “abstract” value in $W$.)

These observations imply that $W$ is a final object in $K$, that is: an abstract data type is a
final object in the category of its representations (where of course, “abstract data type” means abstract in the sense of [18]).

A second argument for this thesis (particularly in support of the uniqueness condition) may be made as follows: the correctness of a data type representation is proved (in [18]) relative to a particular abstraction function. Thus an implementation is a pair \((A, \mathcal{A})\) where \(A\) is a \(T\)-algebra and \(\mathcal{A}\) is an abstraction map \(A \rightarrow W\). This makes \(K\) a “comma category” whose objects are pairs \((A, \mathcal{A})\) as sketched above and whose morphisms \((A, \mathcal{A}) \rightarrow (B, \mathcal{B})\) are \(T\text{-Alg}\) morphisms \(h: A \rightarrow B\) such that the diagram

\[
\begin{aligned}
A & \xrightarrow{h} R \\
\mathcal{A} & \downarrow \downarrow \mathcal{B} \\
W &
\end{aligned}
\]

commutes. Then \((W, 1)\) is a final object, since the diagram

\[
\begin{aligned}
A & \xrightarrow{h} W \\
\mathcal{A} & \downarrow 1 \\
W &
\end{aligned}
\]

commutes iff \(h = \mathcal{A}\).

4. DATA TYPE EXTENSIONS

Guttag [16] has suggested concentration on the issue of data type extensions—that is, the process of adding new types to existing type structures. In the example of Section 1, we extended \(T_{1,+}\) to \(T_{ARR}\). This extension is presented by adding new generators and relations to the generators and relations in the presentation of \(T_{1,+}\). A presentation of a data type extension, then, might be a 4-tuple \((\Omega_0, A_0, \Omega_1, A_1)\) where \((\Omega_0, A_0)\) is a presentation of a base theory \(T_0\) (e.g., \(T_{1,+}\)), and \(\Omega_1\) and \(A_1\) are new generators and relations to be “added”. The theory \(T_1\) of old and new data types is (roughly) \(F_{\Omega_0 \cup \Omega_1}(A_0 \cup A_1)\).

What we are trying to present is a functor \(T_0 \rightarrow T_1\); that is, we are trying to specify both the new theory \(T_1\) and its relation to the base theory \(T_0\). What requirements should be placed on this functor? Clearly, it should be product-preserving. One might require sorts of \(T_0\) to be mapped to sorts in \(T_1\), but for our purposes this is unnecessary. One would be upset if the additional identities in \(T_1\) caused values in \(T_0\) to merge (e.g., if in \(T_{ARR}\) we could conclude that \(n_2 = n_0\)). For this purpose we could ask that the functor be faithful.

Guttag [16] proposed a new condition for data type extensions. He suggested that a presentation of a data type extension was “sufficiently-complete” iff any term in \(F_{\Omega_0 \cup \Omega_1}(A, a)\), where \(a\) is a sort in \(T_0\), is reducible via identities in \(A_0 \cup A_1\) to a term in \(T_0\). The appropriate condition on the functor is \(A\)-fullness, which is defined as follows:
DEFINITION. Let \( T_0 \) be an \( S \)-sorted theory, and let \( C \) be any category. A functor \( i: T_0 \to C \) is \( \Lambda \)-\textbf{full} (respectively \( \Lambda \)-\textbf{faithful}) iff for every \( a \in S \), the function \( T_0(\Lambda, a) \to C(i(\Lambda), i(a)) \) given by \( f \mapsto i(f) \) is surjective (resp., injective).

If a data type extension functor is \( \Lambda \)-full, then no "new" values of the old types will be present in the initial algebra of \( T_1 \). Note that the functor \( T_{1,+} \to T_{ARR} \) is \( \Lambda \)-full but not full; the term
\[
\text{val}[\text{alt}[\emptyset, x_1, x_2], x_3]
\]
is not equivalent to any morphism of \( T_{1,+} \).

DEFINITION. A \textit{data type extension} is a functor \( i: T_0 \to T_1 \) where \( T_0 \) and \( T_1 \) are algebraic theories, and \( i \) is product-preserving, \( \Lambda \)-full, and \( \Lambda \)-faithful.

Given a data type extension \( i \), what is its category of representations? Clearly it should be a subcategory of \( T_1\text{-Alg} \). Thus a typical representation of \( i \) is shown in Figure 4.1.

![Figure 4.1](image)

Having imposed the \( \Lambda \)-faithfulness condition on \( i \) to ensure that \( T_1 \) does not merge values in \( T_0 \), we would not like this information to be lost by \( A \). Therefore we require \( A \circ i \) to be \( \Lambda \)-faithful. For example, if \( i \) is the data type extension \( T_{1,+} \to T_{ARR} \), the \( T_{ARR} \)-algebra \( W: T_{ARR} \to \text{Sets} \) (defined in Section 3) is not \( \Lambda \)-faithful (it merges just those array values in \( T_{ARR}(\Lambda, a) \) which we felt deserved merging), but \( W \circ i \) is \( \Lambda \)-faithful (the integers don't get merged).

We impose a second condition on implementations: a "reachability" condition, which means that an implementation of \( i \) has no values except required by \( T_1 \).

DEFINITION. If \( i: T_0 \to T_1 \) is a data type extension, the category \( K_i \) of implementations of \( i \) is the full subcategory of \( T_1\text{-Alg} \) consisting of product preserving functors \( A: T_1 \to \text{Sets} \) such that

1. \( A \circ i \) is \( \Lambda \)-faithful and
2. for each object \( w \) of \( T_1 \), the map \( \eta_w^A: T_1(\Lambda, w) \to A(w) \) given by \( f \mapsto Af(\ ) \), is surjective.

Condition (2) is worthy of more explanation for the noninitiate. If \( f \in T_1(\Lambda, w) \), then \( Af \in \text{Sets}(A(\Lambda), A(w)) \). Thus \( Af \) is a function of no arguments, yielding a value in \( A(w) \). Thus \( Af(\ ) \), being the application of \( Af \) to a string of no arguments, evaluates to this value. Another condition equivalent to condition (2) is that for each \( w \), the map \( T_1(\Lambda, w) \to \text{Sets}(A(\Lambda), A(w)) \) given by \( f \mapsto Af \) is surjective.

\(^9\) Since \( i \) itself uniquely determines \( T_0 \) and \( T_1 \), we say "implementation of \( i \)" rather than "implementation of \( T_1 \) relative to \( T_0 \) via \( i \)" or the like.
PROPOSITION 1. Let $T$ be any theory. A $T$-algebra $A: T \rightarrow \text{Sets}$ is initial iff for every object $w$ of $T$, $\eta_w^A$ is bijective.

Proof. In the canonical initial $T$-algebra $C$, $\eta_w^C(f) = f$, so $\eta_w^C$ is bijective. Since all initial $T$-algebras are isomorphic, $\eta_w^B$ is bijective for every initial $T$-algebra $B$.

If $A$ is any $T$-algebra and $\eta_w^A$ is bijective, $\eta_w^C \circ (\eta_w^A)^{-1}$ is an isomorphism between $A$ and the canonical initial $T$-algebra $C$.

PROPOSITION 2. If $i: T_0 \rightarrow T_1$ is a data type extension, and $A$ is an object of $K_i$, then $A \circ i: T_0 \rightarrow \text{Sets}$ is an initial $T_0$-algebra.

Proof. $A \circ i$ is $A$-faithful, so for each object $w$ of $T_0$, $\eta_w^{A \circ i}$ is injective. $\eta_w^A$ is surjective, so $\eta_w^{A \circ i} = \eta_w^A \circ i$ is surjective as well. Hence, by Proposition 1, $A \circ i$ is initial.

This proposition establishes that in the terminology of [10] the sorts of $T_0$ are “protected” in $T_1$. The map $A \mapsto A \circ i$ (“composition with $i$”) is the forgetful functor $T_1\text{-Alg} \rightarrow T_0\text{-Alg}$ mentioned in [10, Definition 9].

PROPOSITION 3. Let $i: T_0 \rightarrow T_1$ be a theory-functor. Then $i$ is a data type extension iff for every object $A$ of $K_i$, $A \circ i$ is an initial $T_0$-algebra.

Proof. The “only-if” was shown in the previous proposition. For the reverse direction, let $A$ be an initial $T_1$-algebra. $\eta_w^{A \circ i}$ is a bijection by hypothesis, and $\eta_w^A$ is a bijection by Proposition 1. For any $f \in T_0(A, w)$, $\eta_w^{A \circ i}(f) = \eta_w^A(i(f))$. Hence $i(f) = (\eta_w^A)^{-1} \circ \eta_w^{A \circ i}(f)$. So $i$ restricted to $T_0(A, w)$ is a bijection, and $i$ is $A$-full and $A$-faithful.

5. RESULTS

We have now returned to the situation we found in Section 3: we have a category $K_i$ of data type representations. Can we find an abstract data type in $K_i$? Our main theorem gives an affirmative answer:

MAIN THEOREM. If $i: T_0 \rightarrow T_1$ is a data type extension, then $K_i$ has a final object.

We begin with a characterization of the objects of $K_i$, given by Theorem 1.

DEFINITION. Let $i: T_0 \rightarrow T_1$ be a data type extension. A congruence $Q$ on $T_1$ is $i$-faithful iff the composite $T_0 \rightarrow i^* T_1 \rightarrow T_1/Q$ is $A$-faithful.

LEMMA 1. Let $i: T_0 \rightarrow T_1$ be a data type extension. A congruence $Q$ on $T_1$ is $i$-faithful iff for every pair of morphisms $f, g \in T_0$ with domain $A$, if $(i(f), i(g)) \in Q$, then $f = g$.

THEOREM 1. Let $i: T_0 \rightarrow T_1$ be a data type extension. A product-preserving functor
$A: T_1 \rightarrow \text{Sets}$ is an object of $K_i$ iff there exists an $i$-faithful congruence $Q$ on $T_1$ such that the diagram

\[
\begin{array}{ccc}
T_1 & \rightarrow & T_1/Q \\
\downarrow & & \downarrow \\
\text{Sets} & & \text{Sets}
\end{array}
\]

commutes, where $B$ is an initial algebra of $T_1/Q$.

**Proof.** $(\Rightarrow)$: Given an $i$-faithful congruence $Q$, we merely observe that $B$ is faithful.

$(\Leftarrow)$: Given an object $A$ of $K_1$, let $Q(w, v) = \{(f, g) \mid f, g \in T_1(w, v) \text{ and } A(f) = A(g)\}$. $Q$ is the usual kernel congruence. To show that $Q$ is $i$-faithful, let $f, g \in T_1(A, w)$ for some $w$, and let $(i(f), i(g)) \in Q(A, w)$. Then $A(i(f)) = A(i(g))$. Since $A \circ i$ is $A$-faithful, we conclude $f = g$. Hence $Q$ is $i$-faithful by Lemma 1. Let $[f]$ denote the equivalence class of $f$ modulo $Q$.

Let $j$ denote the functor $T_1 \rightarrow T_1/Q$. Let $B$ denote the $T_1/Q$-algebra given by

\[
\begin{align*}
B(w) &= A(w) \\
B([f]) &= A(f) & f & \in T_1(w, v)
\end{align*}
\]

The second portion of the definition is independent of representatives by the construction of $Q$. Hence $A = B \circ j$.

It remains to show that $B$ is an initial algebra of $T_1/Q$. Let $C$ denote the canonical initial algebra of $T_1/Q$. We claim that $B$ and $C$ are isomorphic in $(T_1/Q)$-$\text{Alg}$. We must give for each object $w$ of $T_1/Q$, a bijective map $\xi_w : B(w) \rightarrow C(w)$ such that $\xi$ is a natural transformation $B \rightarrow C$. In order to do this, let $\eta_w$ denote the surjective map $T_1(A, w) \rightarrow A(w)$ whose existence is guaranteed by the definition of $K_i$. Given $x \in B(w) \leftarrow A(w)$, let $\eta_w^{-1}(x)$ denote any $f \in T_1(A, w)$ such that $A(f) = x$. Now let $\xi_w(x) = j(\eta_w^{-1}(x))$. The value of $\xi_w(x)$ is independent of the choice of $\eta_w^{-1}$ by construction of $Q$. It is easy to see that $\xi_w$ is bijective. To show that $\xi$ is a natural transformation, we must show that for any $[f] \in T_1/Q(w, v)$, $C([f]) \circ \xi_w = \xi_w \circ B([f])$. Let $x \in B(w) \leftarrow A(w)$. Then

\[
\begin{align*}
C([f]) \circ \xi_w(x) &= (C([f]))(j(g)) & \text{where } \eta_w(g) &= x \\
&= (C([f]))(g) \\
&= [gf]
\end{align*}
\]

and

\[
\begin{align*}
\xi_w \circ B([f])(x) &= \xi_w \circ A(f)(x) \\
&= \xi_w \circ A(f) \circ A(g)( \ ) & \text{since } x = \eta_w(g) = (Ag)( \ ) \\
&= \xi_w \circ A(gf)( \ ) \\
&= \xi_w \circ \eta_w(gf) \\
&= j \circ \eta_w^{-1} \circ \eta_w(gf) \\
&= j(gf) \\
&= [gf].
\end{align*}
\]
Lemma 2. Let \( i: T_0 \rightarrow T_1 \) be a data type extension, let \( A, B \) be objects of \( K_i \), and for each object \( w \) of \( T_1 \), let \( \xi_w \) be a map \( A(w) \rightarrow B(w) \). Then \( \xi \) is a morphism of \( K_i \) iff for each \( w \), \( \xi_w \circ \eta_w^A = \eta_w^B \).

Proof. Let \( \xi \) be any natural transformation from \( A \) to \( B \). Then for any \( f \in T_1(A, w) \) following diagram commutes:

\[
\begin{array}{c}
A(A) \\ \downarrow \quad \downarrow \\
B(A) \\
\end{array}
\]

Notice that since \( A(A) \) and \( B(A) \) are singleton sets, the topmost arrow is unique. Chasing the unique member of \( A(A) \) around the diagram, we conclude that \( \xi_w(\eta_w^A(f)) = \eta_w^B(f) \) for every \( f \). Hence \( \xi_w \circ \eta_w^A = \eta_w^B \).

For the "if" portion, assume that for each \( w \), \( \xi_w \circ \eta_w^A = \eta_w^B \). By surjectivity of \( \eta_w^A \), it will suffice to show \( \xi_w \circ Af \circ \eta_w^A = Bf \circ \xi_w \circ \eta_w^A \) (See Figure 5.1).

We chase an element \( g \) of \( T(A, w) \) around the diagram as follows:

\[
\begin{align*}
\xi_w \circ Af \circ \eta_w^A(g) &= \xi_w \circ Af \circ Ag(\ ) \quad \text{(Def'n of } \eta_w^A) \\
&= \xi_w \circ A(fg)(\ ) \quad \text{(} A \text{ is a functor)} \\
&= \xi_w \circ \eta_w^A(fg) \quad \text{(Def'n of } \eta_w^A) \\
&= \eta_w^B(fg) \quad \text{(} \xi_w \circ \eta_w^A = \eta_w^B \text{)} \\
&= B(fg)(\ ) \quad \text{(Def'n of } \eta_w^B) \\
&= Bf \circ Bg(\ ) \quad \text{(} B \text{ is a functor)} \\
&= Bf \circ \eta_w^B(g) \quad \text{(Def'n of } \eta_w^B) \\
&= Bf \circ \xi_w \circ \eta_w^A(g) \quad \text{(} \eta_w^B = \xi_w \circ \eta_w^A \text{).}
\end{align*}
\]

So \( \xi \) is a natural transformation.

Lemma 3. If \( i: T_0 \rightarrow T_1 \) is a data type extension, and \( A, B \) are objects of \( K_i \), then there is at most one morphism from \( A \) to \( B \) in \( K_i \).
Proof. Let \( \xi, \xi' \) be two natural transformations from \( A \) to \( B \). Then \( \xi_w \circ \eta^A_w = \eta^B_w = \xi'_w \circ \eta^A_w \). By surjectivity of \( \eta^A_w \), \( \xi_w = \xi'_w \).

We now need a theorem about least upper bounds of sets of congruences.

**Theorem 2.** Let \( T \) be an \( S \)-sorted theory. Let the congruences on \( T \) be ordered by inclusion. Let \( \mathcal{Q} \) be any set of congruences on \( T \). Then \( \mathcal{Q} \) has a least upper bound, denoted \( \cup \mathcal{Q} \), characterized as follows: If \( f, g \in T(w, v) \), then \( (f, g) \in \cup \mathcal{Q} \) iff there exists \( f_0, \ldots, f_n \in T(w, v) \) such that

(i) \( f_0 = f \)

(ii) \( f_n = g \)

(iii) for each \( i, 0 < i < n \), there exists \( Q \in \mathcal{Q} \) such that \( (f_i, f_{i+1}) \in Q(w, v) \).

**Proof.** Let \( Z(w, v) = \{(f, g) \mid (f, g) \in Q(w, v) \text{ for some } Q \in \mathcal{Q}\} \). We claim that \( Z \) is the desired least upper bound of \( \mathcal{Q} \). If \( Q \in \mathcal{Q} \), then \( Q \subseteq Z \subseteq E_Z \). If \( R \) is a congruence and \( Q \subseteq R \) for each \( Q \in \mathcal{Q} \), then \( Z \subseteq R \), so \( E_Z \subseteq E_R = R \). So \( E_Z \) is the least upper bound.

It remains to show that \( (f, g) \in E_Z \) iff there exists a sequence \( f_0, \ldots, f_n \) as specified. If the sequence exists, then \( (f, g) \in E_Z \) by repeated application of transitivity (rule ET). We will next show that if \( (f, g) \in E_Z \), then the sequence exists. The proof is by induction on derivations in the formal system \( E_Z \). If \( (f, g) \in E_Z \) via an axiom, then it is easy to see that the required sequence exists. For each rule we will have an induction step of the form: "if sequences exist for the hypotheses of the rule, then a sequence exists for the conclusion of the rule." (In the following steps, we write "\( f_0, \ldots, f_n \) is a sequence for \( (f, g) \)" to mean \( f_0, \ldots, f_n \) satisfies condition (iii) of the theorem, \( f_0 = f \), and \( f_n = g \).) Names of quantities in the rules are taken from Section 2.

(ES): If \( f_0, f_1, \ldots, f_n \) is a sequence for \( (f, g) \), then \( f_n, f_{n-1}, \ldots, f_0 \) is a sequence for \( (g, f) \).

(ET): If \( f_0, \ldots, f_n \) is a sequence for \( (f, g) \), and \( g_0, \ldots, g_m \) is a sequence for \( (g, h) \), then \( f_0, \ldots, f_n, g_0, g_1, \ldots, g_m \) is a sequence for \( (f, h) \).

(EC): if \( f_0, \ldots, f_n \) is a sequence for \( (f, f') \), then for each \( i, 0 \leq i < n \), there exists \( Q \in \mathcal{Q} \) such that \( (f_i, f_{i+1}) \in Q \). Since \( Q \) is a congruence, \( (gf, h, gf, h) \in Q \) as well. Hence \( gf, h, gf, h \) is a sequence for \( (gf, h, gf, h) \).

(EP): For \( 1 \leq i \leq n \), let \( f_0, \ldots, f_i \) be a sequence for \( (f_i, f_i) \). Let \( P(k) = \sum_{i=1}^{k-1} p_i \). We construct a sequence \( g_j \) \((0 \leq j \leq P(n + 1))\) for \( (f_i, \ldots, f_n) \) by specifying the projections of the \( g_j \):

\[
E_k g_j = \begin{cases} 
  f_k & \text{if } j < P(k) \\
  f_{ki} & \text{if } P(k) \leq j < P(k + 1); \quad i = j - P(k) \\
  f'_{ki} & \text{if } j \leq P(k + 1).
\end{cases}
\]

The effect of this construction is to create a sequence which changes one component at a time:
In the argument for rule $EC$, for each step in the sequence, $(g_1, g_{j+1})$ is in some $Q \in \mathcal{Q}$ because the pair of components which change is in some congruence $Q$, so the whole step is in $Q$ by applying rule $EP$ for $Q$. Verification of the details is left to the diligent reader. 

Note: Theorem 2 is a variation of a theorem well-known for single-sorted algebras [15, Lemma 10.2]. An alternate proof could be obtained by proving the theorem for many-sorted algebras in general [2], and then observing that an $S$-sorted theory is itself an $S^* \times S^*$-sorted algebra [1].

**Theorem 3.** Let $i: T_0 \to T_1$ be a data type extension, and let $\mathcal{Q}$ be a set of $i$-faithful congruences on $T_1$. Then $\bigvee \mathcal{Q}$ is $i$-faithful.

**Proof.** Let $f, g \in T_0(A, w)$, with $(i(f), i(g)) \in \bigvee \mathcal{Q}$. Then by Theorem 2 there exist $f_0, \ldots, f_n \in T_1(A, w)$ such that $f_0 = i(f), f_n = i(g)$, and for each $j, 0 \leq j < n$, $(f_j, f_{j+1}) \in Q$ for some $Q \in \mathcal{Q}$. Since $i$ is a data type extension, it is $A$-full, so for each $f_j$ there exists $g_j \in T_0(A, w)$ such that $f_j = i(g_j)$. By Lemma 1, $g_j = g_{j+1}$. Hence $f_j = f_{j+1}$, and $f_0 = f_n$. Since $i$ is $A$-faithful, $f = g$. By Lemma 1, this establishes that $\bigvee \mathcal{Q}$ is $i$-faithful. 

We are now ready to prove the main theorem.

**Proof of the main theorem.** Let $\mathcal{Q}$ be the set of all $i$-faithful congruences on $T_1$. Let $\hat{Q} = \bigvee \mathcal{Q}$, let $\hat{T} = T_1/\hat{Q}$, and let $j$ be the quotient functor $T_1 \to \hat{T}$. Let $W$ be the canonical initial algebra of $\hat{T}$, and let $C = W \circ j: T_1 \to \text{Sets}$. We claim that $C$ is a final object of $K_i$. By Theorem 3, $\hat{Q}$ is $i$-faithful, so by Theorem 1, $C$ is an object of $K_i$.

Now let $A$ be any object of $K_i$. By Theorem 1, there exists an $i$-faithful congruence $Q$ on $T_1$, with quotient functor $j: T_1 \to T_1/Q$, and an initial algebra $B$ of $T_1/Q$ such that $A = B \circ j$. Since $Q$ is $i$-faithful, we have a theory-functor $k: (T_1/Q) \to (T_1/Q) = \hat{T}$ such that $j = k \circ j$.

By Lemma 2, we need only show that for each object $w$ of $T_1$, $\eta_w^C$ factors through $\eta_w^A$. We claim that $\eta_w^C = k \circ (\eta_w^B)^{-1} \circ \eta_w^A$. Since $B$ is an initial $T_1/Q$-algebra, $\eta_w^B$ is a bijection, so $(\eta_w^B)^{-1}$ is well defined. So, if $f \in T_1(A, w)$, then

\[
k \circ (\eta_w^B)^{-1} \circ \eta_w^A(f) = k \circ (\eta_w^B)^{-1}(Af) = k \circ (\eta_w^B)^{-1}(Af) = k \circ j(f) = j(f) = \eta_w^C(f)
\]

As justification for the step marked (*), we calculate:

\[
\eta_w^B(j(f)) = B(j(f))(A) = Af(\ )
\]

Hence $\xi_w = k \circ (\eta_w^B)^{-1}$ is the required natural transformation. Uniqueness is guaranteed by Lemma 3.
6. Example

In this section we will complete our consideration of the array example.

Proposition 3. Let $i$ be the data type extension $T_{I,+} \rightarrow T_{ARR}$. Then the $T_{ARR}$-algebra $W$, defined in Section 3, is a final object of $K_i$.

Proof. First, it is straightforward to check that $W$ is an object of $K_i$. Furthermore, if $f \in T_{ARR}(A, a)$, it is easy to show from the axioms for $T_{ARR}$ that the partial function $W(f)$ is defined at just those integers $j$ such that in $T_{ARR}$, $\text{val}[f, n_j] = n_k$ for some $k$, and that for any other $j$, $\text{val}[f, n_j] = \text{undefined}$.

Now, let $A$ be any object of $K_i$. By Lemma 2, to get a morphism $\xi: A \rightarrow W$, we must show that $\eta^W$ factors through $\eta^A$. Since the $\eta$'s are in the category of sets, we need only show that for all $f, g \in T_A(A, a)$, if $\eta^A(f) = \eta^A(g)$, then $\eta^W(f) = \eta^W(g)$. Because $A$ and $W$ preserve products, it is enough to prove this for the case where $a$ is a single sort. The integer sorts of all the algebras in $K_i$ are isomorphic (they are initial algebras of $T_{I,+}$, by Proposition 2), so the only interesting case is where $a = (a)$ (the array sort).

In the following, we will write $\eta^x$ for $\eta_a^x$.

Let $f, g \in T_{ARR}(A, a)$. We still show that if $\eta^A(f) \neq \eta^A(g)$, and $\eta^A(f) = \eta^A(g)$, then $A \circ i$ is not $A$-faithful. There are two ways in which $\eta^A(f) = \eta^A(g)$ could be unequal. They may have different domains, or they may have different values at some point of their domain.

If they have different domains, then without loss of generality there exists some $n_p \in T_{ARR}(A, i)$ such that in $T_{ARR}$, $\text{val}[f, n_p] = \text{undefined}$ and $\text{val}[g, n_p] = n_k$ for some $k$. If $\eta^A(f) = \eta^A(g)$, then $\eta^A(\text{undefined}) = A(\text{val})[\eta^A(f), \eta^A(n_p)] = A(\text{val})[\eta^A(g), \eta^A(n_p)] = \eta^A(n_k)$. So $A \circ i$ is not $A$-faithful.

Similarly, if they are unequal at some point in their domain, there exists some $n_p \in T_{ARR}(A, i)$ such that in $T_{ARR}$, $\text{val}[f, n_p] = n_j$ and $\text{val}[g, n_p] = n_k$, for $k \neq j$. If $\eta^A(f) = \eta^A(g)$, it then follows that $\eta^A(n_j) = \eta^A(n_k)$, again showing that $A \circ i$ is not $A$-faithful.

We have considered initial and final algebras in $K_i$. Are there any interesting intermediate cases? The answer is yes. Add, for example, to the axioms for $T_{ARR}$ the following axioms:

- $\text{alt}[\text{alt}[x, n_k, y], n_k, z] = \text{alt}[x, n_k, z]$ \hspace{1cm} $k \in \omega$ \& $k \neq 7$ \& $k \neq 9$
- $\text{alt}[\text{alt}[x, n_k, y], n_p, z] = \text{alt}[\text{alt}[x, n_p, z], n_k, y]$ \hspace{1cm} $k, p \in \omega$ \& $k \neq p$
- $\text{alt}[x, \text{undefined}, y] = x$

The resulting initial algebra suppresses most of the order information but preserves “traces” of all values assigned to location 7 or to location 9. A single trace showing how these assignments were interleaved may be obtained by putting similar restrictions on the second axiom scheme.
7. Final vs. Initial Algebra Semantics

The final algebra constructed in the main theorem is an initial algebra of $T_i/Q$. Why then, do we distinguish "final algebra semantics" from "initial algebra semantics"? The answer lies in the primacy of specification.

We believe that a program should interact with a data type only through its specifications. Thus a specification, which is a formal object in some logical calculus, must present a class of algebras, namely, the class of implementations of the data type.

As we argued in Section 3, a theory of data type representations should make the "true" data type a final object in its category of representations. Methodologically, final algebra semantics is more desirable in this regard.

Methodological considerations aside, it may be that $T_0$ and $T_1$ have tractable presentations, but $T_i/Q$ does not. In our case, $T_0$ and $T_1$ had Church--Rosser presentations, but the obvious presentation of $T_i/Q$ was not Church--Rosser. We leave it open whether there exist finitely presentable $T_0$ and $T_1$ such that $T_i/Q$ is not finitely presentable.

In any case, "final algebra semantics" should be regarded as an extension, rather than a competitor, of initial algebra semantics. If $i$ is the identity functor $T_1 \rightarrow T_1$, then $K_i$ consists entirely of initial $T_1$-algebras. Furthermore, if $T_1$ and $T_i/Q$ happen to be equal, then initial and final algebras coincide again. For example, take $T_0$ to be $T_{1,+}$ as before and let $T_1$ be given by adding a new sort $\omega$ ("string of integers") with $T_1(A, \omega) = \omega^*$ and operations

\[ \text{sel}_k : \omega \rightarrow \omega \quad k \in \omega \]

which select the $k$th integer from a string (or give undefined if the string is too short). Now, any $i$-faithful congruence on $T_1$ must be the equality relation (for if not, assume $\alpha$ and $\beta$ are two distinct congruent strings. Since they are distinct, they must differ at some position (say the $j$th position). Then

\[ n_k = \text{sel}_k \alpha = \text{sel}_k \beta = n_p \]

for $k \neq p$, violating $i$-faithfulness.) So again $K_i$ consists of initial $T_1$-algebras, but the presentation of $K_i$ by $i$ also specifies the relation of $T_0$ to $T_1$.

8. Concluding Remarks

The situation discussed in this paper is reminiscent of categories of automata, constrained by the requirement that some external behavior be maintained. In our case the "external behavior" is the behavior which is reflected in the sorts of $T_0$; hence the condition that $A \circ i$ be $A$-faithful. One has initial realizations and, if one imposes a reachability condition, one has a minimal realization which is a final object [7]. We, too, have an initial realization (an initial $T_1$-algebra), and a final or minimal realization whose existence is our main result.
Similar remarks are echoed in [6]. Our notion of extension includes all of [10, Def. 9] including enrichment. It also allows the possibility that a single sort in $T_0$ is mapped to a tuple of sorts in $T_1$. We regard this paper as complementary to [10], which seems to be devoted to the problems of specifying $T_1$ (which is no small task!).

Another echo deserving of mention is that of data structure selection and optimization. The initial implementation is a very crude data structure, consisting solely of trees (see e.g. [17]). In our example, arrays are represented as lists of subscript-value pairs (without even deleting updated entries!). By looking at the required updates and addresses, the data structure implementing the data type may be optimized until no redundant information is stored. In our example, arrays turn out to be optimal in this sense.

We leave open the question of formulating a "behavior" functor adjoint to "minimal realization" [7]; such a development might shed some light on the distinction between data structures and data types. Another extension could involve types with type parameters.

REFERENCES