Observation and Abstract Behaviour in Specification and Implementation of State-based Systems

Claus Pahl
School of Computer Applications
Dublin City University
Dublin 9, Ireland
cpahl@compapp.dcu.ie

Abstract

Classical algebraic specification is an accepted framework for specification. A criticism which applies is the fact that it is functional, not based on a notion of state as most software development and implementation languages are. We formalise the idea of a state-based object or abstract machine using algebraic means. In contrast to similar approaches we consider dynamic logic instead of equational logic as the framework for specification and implementation. The advantage is a more expressive language allowing us to specify safety and liveness conditions. It also allows a clearer distinction of functional and state-based parts which require different treatment in order to achieve behavioural abstraction when necessary. We shall in particular focus on abstract behaviour and observation. A behavioural notion of satisfaction for state-elements is needed in order to abstract from irrelevant details of the state realisation.

1 Introduction

Algebraic methods have been widely used in the specification of computer systems. Algebraic specification refers to the use of algebraic semantics and equational reasoning for functional systems [1]. During the last decade, algebraic methods have been used to support the more wide-spread state-based software development. Examples of state-based specification include the pre/post-condition technique originally developed by Hoare [2, 3]. This formal technique has found its way into various software development approaches and languages such as the UML Object Constraint Language [4] or the design-by-contract approach [5]. The pre/post-condition technique can be extended or generalised in various ways [6]. We shall address the more general framework of dynamic logic here, including pre- and postconditions. We shall in particular address semantic foundations for a framework for observational specification of state-based systems in this paper. This shall result in a more flexible specification and implementation framework for state-based systems.

A classical example of the usefulness of observational specification is the realisation of a data type boolean using an implementation of a type integer. The data values needed to interpret terms formulated based on boolean constructors (e.g. true and not) form only a subset of a carrier set for integers. We expect boolean axioms only to hold within this subset. An observability predicate can declare those substructures reachable with boolean constructors as observable. Observation and behaviour oriented extensions to the algebraic specification of data types exist (e.g. [7, 8, 9, 10, 11, 12, 13, 14]), but similar approaches to state-based specification are still lacking. We aim to provide semantical foundations for such a framework of behaviour and observation in the context of state-based systems.

We shall give an outline of our approach here. The classical notion of signatures is extended by introducing hidden, non-observable sorts representing an internal state. Semantic structures interpreting these signatures provide suitable elements. Slightly different techniques apply for defining and reasoning about functional and state-based parts. One
of the most well-known approaches to these problems is the hidden algebra framework developed by the OBJ-group, see e.g. [15, 16]. There, an equational, behaviourally oriented framework for reasoning is established. We follow the hidden algebra approach in its basic idea: we provide signatures with a distinguished state sort, and interpret these in structures with a local state, called objects. An early version of this model has been presented in [17]. Then, we deviate from the path by using concepts from modal logics to introduce a specification and reasoning framework. A dynamic logic will be used. Modal logics, such as dynamic and temporal logics, provide constructs that will allow us to specify safety as well as liveness constraints – a wider range of properties than with equational logics. Safety properties shall guarantee that unwanted things will not occur, and liveness properties should ensure that wanted things will eventually occur. These properties are in particular important in the specification of dynamic and reactive systems. Our particular interest lies in the implementation of specifications. In implementing a specification it has to be shown that the implementation satisfies the abstract requirements of the specifications. In order to allow the simpler equational reasoning to be used, we complement a dynamic logic based implementation with a refinement notion which needs equational reasoning only in proofs. We consider data and operation refinement and show how they relate to the paradigm of observational or behavioural specification, namely, that formulas need only to be true under all possible experiments (or observations).

We start with our semantic model in Section 2. Then, in Section 3, we present the dynamic logic. Section 4 introduces basic concepts for behavioural specifications. In Section 5, we look at data refinement based on behavioural abstraction. In Section 6, the operation refinement is investigated. Proofs, if they are not given here, can be found in [18].

2 A State-based Model

The main building blocks of our computational model – signatures with state, objects as the corresponding models, and an interpretation – shall be formalised in this section. The objective is to formalise the notion of an object - an algebraic entity that encapsulates a local state.

We assume a set of sorts \( \text{Sort} \) including data sorts \( s_1, s_2, \ldots \) and the distinguished sort \( \text{state} \), and a set of identifiers \( \text{Id} \). For each state signature \( \Sigma \) in a class of signatures \( \text{Sig} \) we define:

- a mapping \( \text{sorts} : \text{Sig} \rightarrow \mathcal{P} \text{Sort} \) with \( \text{state} \in \text{sorts}(\Sigma) \),
- a mapping \( \text{attr} : \text{Sig} \rightarrow \mathcal{P} \text{Id} \) with a signature \( \text{sig}(a) = s_1 \times \ldots \times s_n \rightarrow s \) for each attribute \( a \in \text{attr}(\Sigma) \) with \( s_1, \ldots, s_n, s \in (\text{sorts}(\Sigma) - \{\text{state}\})^1 \),
- the state sort \( \text{state} \) as \( \text{attr}(\Sigma) \rightarrow \text{sig}(\text{attr}(\Sigma)) \)
- a mapping \( \text{tran} : \text{Sig} \rightarrow \mathcal{P} \text{Id} \) with a signature \( \text{sig}(p) = \text{state} \times s_1 \times \ldots \times s_m \rightarrow \text{state} \times s \) for each transition \( p \in \text{tran}(\Sigma) \).

Basic sorts – not including \( \text{state} \) – are typically data types used in the definition of attributes. We assume a built-in sort \( \text{bool} \) with the usual constants and functions. The set of attributes, variables (nullary functions) and other functions, forms an explicit, but still abstract state. The state itself is represented as an assignment, associating attribute identifiers with functions. These functions are the only means to inspect the state content. Transitions are parameterised state transformers, transforming one state into another, also yielding a functional result value.

Example 2.1 Sample sorts in a stack data type are the stack content stack – which shall represent the state-sort – and the data type elem. The following \( \Sigma \) is a signature:

\[
\begin{align*}
\text{sorts}(\Sigma) &= \{\text{state}, \text{elem}, \text{bool}\} \\
\text{attr}(\Sigma) &= \{\text{top} : \rightarrow \text{elem}, \text{is\_full} : \rightarrow \text{bool}\} \\
\text{tran}(\Sigma) &= \{\text{push} : \text{state} \times \text{elem} \rightarrow \text{state}, \text{pop} : \text{state} \rightarrow \text{state}\}
\end{align*}
\]

\(^1a \) can be a nullary symbol, then called a state variable.
We can now capture the idea of a state-based computational model precisely. An algebraic structure is called a **state signature** \(\Sigma\) including variables. A term of sort \(s\) is a **term** of the signature. In many-sorted signatures, a sort will be associated with each term. Let \(\Sigma\) be a state signature. A signature induces a set of syntactically correct expressions, constructed by free variables and the operation names \(\text{op}\) and \(\text{attr}\). A **signature morphism** \(\sigma\) is a signature morphism where \(\sigma_{\text{sort}}(\text{state})\) is the identity, and for any operation \(\text{op} \in \text{attr}(\Sigma) \cup \text{tran}(\Sigma)\) with \(\sigma(\text{op}) = a_1 \times \ldots \times a_i \rightarrow b_1 \times \ldots \times b_j\) we require \(\sigma_{\text{op}}(\text{op}) = \sigma_{\text{sort}}(a_1) \times \ldots \times \sigma_{\text{sort}}(a_i) \rightarrow \sigma_{\text{sort}}(b_1) \times \ldots \times \sigma_{\text{sort}}(b_j)\) for appropriate \(i, j\), where \(\sigma_{\text{sort}}\) is a mapping on sorts and \(\sigma_{\text{op}}\) a mapping on attributes and transition identifiers. A subsignature \(\Sigma'\) can be embedded into a signature \(\Sigma\) by an embedding signature morphism.

**Example 2.2** Consider the signature \(\Sigma'\):

\[
\begin{align*}
\text{sorts}(\Sigma') & = \{\text{state}, \text{elem}\} \\
\text{attr}(\Sigma') & = \{\text{top} : \text{elem}\} \\
\text{tran}(\Sigma') & = \{\text{push} : \text{state} \times \text{elem} \rightarrow \text{state}, \text{pop} : \text{state} \rightarrow \text{state}\}
\end{align*}
\]

\(\Sigma'\) is a subsignature of \(\Sigma\) – see Example 2.1. The embedding of \(\Sigma'\) into \(\Sigma\) is a signature morphism.

A signature induces a set of syntactically correct expressions, constructed by free variables and the operation names of the signature. In many-sorted signatures, a sort will be associated with each term. Let \(X = (X_s)_{s \in \text{sort}(\Sigma)}\) be a \(\text{sort}(\Sigma)\)-sorted set of free variables. The functions \(\pi_1\) and \(\pi_2\) are projections onto the first and second component of their argument, respectively.

- **Sigma-terms of data sort** \(s \in \text{sorts}(\Sigma)\) over \(X_s\) are variables \(x \in X_s\); every \(a(a_1, \ldots, a_n)\) for \(a \in \text{attr}(\Sigma)\) with \(a : s_1 \times \ldots \times s_n \rightarrow s\) and \(a_i : s_i\); every projection \(\pi_2(p(st, a_1, \ldots, a_m))\) for \(p \in \text{tran}(\Sigma)\) with \(p : \text{state} \times s_1 \times \ldots \times s_m \rightarrow \text{state} \times s, st : \text{state}\) and \(a_i : s_i\).

- **Sigma-terms of state sort** \(\text{state} \in \text{sorts}(\Sigma)\) over \(X_s\) are identifiers \(st\) to denote the state; and every projection \(\pi_1(p(st, a_1, \ldots, a_m))\) for \(p \in \text{tran}(\Sigma)\) with \(p : \text{state} \times s_1 \times \ldots \times s_m \rightarrow \text{state} \times s, st : \text{state}\) and \(a_i : s_i\).

**Example 2.3** A \(\Sigma\)-term of data sort \(\text{bool}\) is \(\text{is}_\text{full()}\). A \(\Sigma\)-terms of state sort \(\text{state}\) are \(\text{pop}(st)\) or \(\text{push}(st, e)\).

With \(T(\Sigma)_s\) we denote the set of ground \(\Sigma\)-terms of sort \(s\). With \(T(\Sigma, X)_s\) we denote the set of \(\Sigma\)-terms of sort \(s\) including variables. A term of sort \(\text{state}\) is also called a **command**.

We can now capture the idea of a state-based computational model precisely. An algebraic structure is called a **state object** for a state signature \(\Sigma\), if it has

- a carrier set \(S\), including an undefinedness symbol \(\perp\), for each sort \(s\).
- a function of type \(S_1 \times \ldots \times S_n \rightarrow S\) for each attribute with signature \(s_1 \times \ldots \times s_n \rightarrow s\).
- a carrier set \(\text{State}\) for sort \(\text{state}\) containing total assignments \(Id : F\) if \(F\) is the set of functions that match signatures of attributes.
- a designated initial state \(ST_{\text{init}} \in \text{State}\).
• a function of type \( (State \times S_1 \times \ldots \times S_m) \rightarrow (State \times S) \) for each transition symbol in \( tran(\Sigma) \) with the corresponding signature.

\( \Sigma \)-objects can be seen as abstract machines with operations as instructions. Initial states (and the ability to specify properties of the initial state) are essential in the context of dynamic or reactive systems. Underlying each of these models, we assume an equationally specified algebra, realising the data type(s) used in the definition of the attributes. These algebras are static, i.e. their functions can not be changed by state transitions. \( \Sigma \)-objects are extensions of these algebras. The state can also be seen as an algebra consisting of carrier sets for data sorts and functions for the attributes. The \( \Sigma^* -\text{state algebra} \) for a \( \Sigma \)-object shall be defined as the algebra consisting of carrier sets for the data sorts (excluding \( state \)) and functions for the attributes \( attr(\Sigma) \). The signature \( \Sigma^* \) consists of the data sorts and attribute signatures of \( \Sigma \). This shows that \( \Sigma \)-objects fit into the states-as-algebras paradigm.

A **hidden algebra** is an algebra which satisfies a hidden specification based on a hidden signature [15]. A hidden algebra should encapsulate an algebra as a substructure (reduct) which represents the data part. Our objects are hidden algebras in this sense.

A \( \Sigma \)-object involves bindings: functions are bound to attribute and transition identifiers. A state \( ST \) of sort \( state \) can be modified by transitions. We need a substitution mechanism on states. The expression \( \text{substitute}(ST, x \mapsto v) \) substitutes in the mapping \( ST \) the former binding for an identifier \( x \) by a binding of \( x \) to the value \( val \):

\[
\text{substitute}(ST, x \mapsto val)(z) = \begin{cases} 
ST(z) & \text{for } z \neq x \\
val & \text{for } z = x 
\end{cases}
\]

Domain and range of \( ST \) shall be disjoint. Then, \( x \) cannot occur in \( val \).

A mapping \( v \) from identifiers to semantical entities is called a **valuation**; its inductively defined extension \( v^* \) for arbitrary terms is called an **interpretation**. Each term depends on the current state. Let \( ST : state \) be a state, i.e., an assignment\(^2\) of functions to identifiers.

\[
v(ST, x) := ST(x)
\]

The value of an identifier \( x \) is stored in \( ST \) – remember the definition of sort state: \( state = attr(\Sigma) \rightarrow sig(attr(\Sigma)) \), i.e., that the state associates attributes and functions.

\[
v^*(ST, a(a_1, \ldots, a_n)) := \\
v^*(ST, p(ST, a_1, \ldots, a_m)) := \\
\begin{array}{l}
ST(a(a_1, \ldots, a_n)) \\
\left[p\left[(ST, v^*(ST, a_1), \ldots, v^*(ST, a_m))\right]\right]
\end{array}
\]

The definition of an attribute \( a \) can be modified, thus we have to reinterpret \( a \) in each current state \( ST \). Transitions \( p \) might modify the definition of attributes. The brackets \( [\ ] \) are used to denote the function which realises a transition symbol.

Finally, we introduce a **definedness** predicate: \( D(ST, t) = true \), if \( v^*(ST, t) \neq \bot \) for some state \( ST \). Termination of operations will be expressed using the definedness predicate.

**Example 2.4** A \( \Sigma \)-object \( A \) may contain an attribute top. state is the local state variable and top is an attribute based on the content of state. States can be transformed into another by push and pop. push is a transition and, therefore, changes the state. By executing push(st,e) the state is changed by substituting the binding of top through assigning a new value \( (v^*(ST, e)) \) to the attribute identifier top. The effect of push can be observed by applying top.

\(^2\)We use the term assignment for the state and valuation for identifiers of the signature in general.
3 State Transition Logic

Modal logics assume structures with a notion of state or time [19, 20]. The presented model of $\Sigma$-objects is such a structure. Dynamic logic, a particular modal logic, is a first-order predicate logic with a notion of state, which generalises Hoare-logic. Based on these logics, we will introduce a simple dynamic logic. A proof system is not presented here, see [18] for a sound and relatively complete one for a general dynamic logic and [21] for one geared towards the use in refinement calculi.

The state transition formula $\phi \rightarrow [P] \psi$ – where $\phi$ is a precondition, $P$ is a command (a term of sort state) and $\psi$ is a postcondition – describes a state transition axiomatically. The operator $[.]$ is called the modal box- or always-operator. $\phi$ and $\psi$ describe properties of attributes by classical first-order formulas. The formula $\phi \rightarrow [P] \psi$ holds in a current state $ST$ in a $\Sigma$-object $A$, if in case $\phi$ holds in $ST$ and, if $P$ terminates, then $\psi$ has to hold in the following state. The second modal operator is the so-called diamond- or eventually-operator. This operator allows us to specify liveness properties. The formula $<P> \phi$ expresses that $P$ has a terminating path after which $\phi$ holds. In case of deterministic programs, it says that $P$ will terminate and $\phi$ holds afterwards. For example, $<P> \text{ true}$ is a termination assertion. These constructs will be defined formally later on.

Example 3.1 The effect of the stack transition push can be specified by

$$\neg \text{is\_full()} \rightarrow [\text{push(st,e)}] \text{ top()}=e$$

If the precondition – the stack is not full – is satisfied, then executing $\text{push(st,e)}$ on stack $st$ has the effect that $e$ is the new top element.

$$\text{is\_full()} \rightarrow <\text{push(st,e)}> \text{ true}$$

push should terminate if the stack is full (whatever the result might be).

Let $\Sigma$ be a state signature and $X$ a sorts($\Sigma$)-sorted set of free variables. A $\Sigma$-equation has the form $t =_s t'$ with $t, t' \in T(\Sigma, X)_s$ for a data sort $s$. A $\Sigma_{state}$-equation has the form $t =_{state} t'$ with $t, t' \in T(\Sigma, X)_{state}$. The set of well-formed formulas $\text{WFF}(\Sigma)$ is the smallest set with the following properties:\footnote{We have omitted classical connectors, such as negation or disjunction, and quantification. They can be defined as usual for first-order predicate logics.}

- all $\Sigma$-equations and $\Sigma_{state}$-equations are in $\text{WFF}(\Sigma)$,
- if $\phi, \psi \in \text{WFF}(\Sigma)$, then $\phi \rightarrow \psi \in \text{WFF}(\Sigma)$,
- if $P$ is a command and $\phi \in \text{WFF}(\Sigma)$, then $[P] \phi \in \text{WFF}(\Sigma)$ and $<P> \phi \in \text{WFF}(\Sigma)$.

We define the diamond operator as follows:

$$<P> \phi := \neg[P] \neg \phi$$

An operator $\text{prev}$ shall also be defined. The operator refers to the value of an argument variable in the previous state of the command $P$ under consideration. Typically, the $\text{prev}$ operator is used to specify the new value of a variable in terms of the old value, e.g. $[P] a() = \text{prev}(a()) + 1$.

A specification $M$ is a pair $M = (\Sigma, E)$ consisting of a state-based signature $\Sigma$ and a set $E$ of well-formed formulas $E \subseteq \text{WFF}(\Sigma)$, including an non-modal formula $\text{Init}$. $\text{Init}$ is a non-modal formula characterising the initial state of a $\Sigma$-object.
Example 3.2 A specification $M = \langle \Sigma, E \rangle$ for a stack consists of:

$$
\Sigma :
$$

- $\text{sorts}(\Sigma) = \{\text{state}, \text{elem}, \text{bool}\}$
- $\text{attr}(\Sigma) = \{\text{top} : \rightarrow \text{elem}, \text{is\_full} : \rightarrow \text{bool}\}$
- $\text{trans}(\Sigma) = \{\text{push} : \text{state} \times \text{elem} \rightarrow \text{state}, \text{pop} : \text{state} \rightarrow \text{state}\}$

$$
E :
$$

- $\neg \text{is\_full()} \rightarrow [\text{push}(\text{st}, e)] \text{top()} = e$
- $\text{top()} = e \rightarrow [\text{push}(\text{st}, e'); \text{pop}()] \text{top()} = e$
- $\text{is\_full()} \rightarrow <\text{push}(\text{st}, e)> \text{true}$

The state is the content of the stack, elem is the element sort. The element $e$ shall be of sort elem and $st$ of sort state. This example assumes that a sequence operator $;$ is defined on transitions. This is not a complete specification of a stack as it, for example, does not specify the effect of pop on empty stacks.

As we have seen in the example, command combinators such as the sequence ‘;’ are useful. We define the combinators $;$ (sequence), $+$ (non-deterministic choice), and $^*$ (iteration):

- $v^*(ST, p; q) := v^*(\pi_1(v^*(ST, p)), q)$
- $v^*(ST, p + q) := v^*(ST, p)$ or $v^*(ST, q)$ – non-deterministically chosen
- $v^*(ST, p^*) := v^*(\pi_1(\ldots(\pi_1(v^*((ST, p)), p), \ldots))$

For two of them, we can find simple axiomatisations:

- $[p; q] \phi \leftrightarrow [p][q] \phi$
- $[p + q] \phi \leftrightarrow [p] \phi \lor [q] \phi$

A notion of satisfaction relates formulas and $\Sigma$-objects. Let $v$ be a valuation based on states and $v^*$ an interpretation. Assume a $\Sigma$-object $A$, a state $ST \in \text{State}$ and a $\Sigma$-formula $\phi$. $A$ satisfies $\phi$ in state $ST$, or $A, ST \models \phi$, is defined by

- $A, ST \models t = s, t'$ if $v^*(ST, t) = v^*(ST, t')$ for a data sort $s$
- $A, ST \models t =_{\text{state}} t'$ if $v^*(\pi_1(v^*(ST, t)), a(a_1, \ldots, a_{n_c})) = v^*(\pi_1(v^*(ST, t')), a(a_1, \ldots, a_{n_c}))$ for all attribute applications
- $A, ST \models \phi \rightarrow \psi$ if $(\neg \phi) \lor \psi$
- $A, ST \models [P] \phi$ if $D(P) \rightarrow A, \pi_1(v^*(ST, P)) \models \phi$ where $D$ is the definedness predicate

Nontermination is denoted by $\bot_{\text{state}}$. Note, that terms of sort state are behaviourally equal. Two terms are only equal if the interpretations of all attribute applications in the subsequent state are equal. This definition was chosen instead of comparing two state structures directly which would have involved non-relevant comparisons.

Example 3.3 The transition push on stack has been specified by

$$
\neg \text{is\_full()} \rightarrow [\text{push}(\text{st}, e)] \text{top()} = e
$$

Only the observable behaviour of push is specified, i.e. if push is executed in a state satisfying $\neg \text{is\_full()}$ and it terminates, then it should terminate in a state whose properties can be inspected by attributes such as top.

\footnote{with the usual definitions for $\neg$ and $\lor$}
The class of all \( \Sigma \)-objects is denoted with \( \text{Obj}(\Sigma) \). The models of a specification \( M = \langle \Sigma, E \rangle \) are denoted by their model class.

\[
\text{mod}(\Sigma, E) := \{ A \in \text{Obj}(\Sigma) \mid A, ST \models \phi \text{ for all } \phi \in E \text{ and } ST \in \text{State} \}
\]

Note that \( E \) contains an initial state condition, i.e. models are only those objects whose designated initial state satisfies that condition. Model construction not requiring closedness with respect to isomorphism is usually called loose. Here, even reachability – the no-junk property – is not required. We call this model semantics very loose.

We will also introduce notions of a model of a transition and a respective satisfaction relation. Let a transition \( p \) in \( M_1 \) be specified by \( \phi_1 \to [p] \psi_1 \) and in \( M_2 \) by \( \phi_2 \to [p] \psi_2 \). Then \( \text{mod}^p_{M_1}(p) \) denotes models of a transition \( p \), i.e. those functions which interpret a transition \( p \) in a model of a specification \( M \). Let \( A \in \text{mod}(M) \), \( [p] \in \text{mod}^p_{M_1}(p) \) and \( p \in \text{tr}\text{an}(\text{sig}(M)) \). Then the satisfaction \( \models_{tr} \) is defined as follows: \( [p] \models_{tr} \phi \) holds, if \( A \models \phi \) holds for all objects \( A \in \text{mod}(M) \) which interpret \( p \) by \( [[p]] \).

**Lemma 3.1** With the previous two definitions, we get the following property for two \( \Sigma \)-formulas \( \phi, \psi \) specifying a transition \( p \) by \( \phi \to [p] \psi \):

\[
[[p]] \in \text{mod}^p_{M_1}(p) \text{ iff } [p] \models_{tr} \phi \to [p] \psi
\]

**Proof:** Obvious. \( \square \)

### 4 Observational Specification

Let us look at notions of behavioural and observational specification now. We briefly present some basic constructs used in the majority of observational algebraic specification, see e.g. [11, 13, 14, 16].

**Definition 4.1** Let \( \Sigma \) be any signature. A \( \Sigma \)-context is a \( \Sigma \)-term, which contains exactly one special variable \( z_s \) for each sort \( s \). The variables \( z_s \) are used as meta-variables for terms of a specific sort.

\[
c[t] := c[z_s / t] \text{ for } t : s
\]

The variable \( z_s \) can be substituted by \( t \) in the term \( c \).

This shall be illustrated by a stack example.

**Example 4.1** \( \text{top}(z) \) shall be a context with special variable \( z \). Then, the equation \( \text{pop}([\text{push}(s, e)]) = s \) does not need to hold, only the application of the observational context to the equation:

\[
\text{top}\text{(pop}(\text{push}(s, e)))) = \text{top}(s)
\]

\( \text{pop}(\text{push}(s, e)) \) and \( s \) substitute \( z \).

In our approach, we would model \( \text{push} \) and \( \text{pop} \) as procedures, since they modify the state, here the stack. \( \text{top} \) would be an attribute.

\[
\text{top}(e) = e \to [\text{seq}(\text{push}(e'), \text{pop}())] \text{ top}(e) = e
\]

This guarantees that only observations on the state using \( \text{top} \) are relevant. This is already behaviourally oriented. We have also defined a behavioural satisfaction for equations on state.

The standard equality can be generalised in observational approaches: two terms shall be considered equal, if they cannot be distinguished through observations. An observational predicate \( {\text{Obs}} \) can be applied to any terms that denote carriers. There is an observational subset for each carrier set. An observational \( \Sigma \)-context is a \( \Sigma \)-context with observational sort.
Example 4.2 For instance $\text{top}(z)$ would be an observational $\Sigma$-context, if the sort of $\text{top}[t]$ for a suitable $t$ is a observational sort.

The notion of a context can also be defined for the state-based setting. A term $c \in T(\Sigma, X)_{s \in S \cup \{\text{state}\}}$ is called context over $\Sigma$, if $c$ contains exactly one $z_s \in Z = \{z_s | s \in S \cup \{\text{state}\}\}$. A context over $z_s$ shall be called observational, if $s$ is not the sort state.

An observational subalgebra $(A, \text{Obs}^A)$ is the subalgebra of $A$ that contains only those elements denotable by ground terms defined as observable. The observational satisfaction $|=_{\text{Obs}}$ is defined by:

$$(A, \text{Obs}^A) =_{\text{Obs}} t = t' \text{ if } A \models c[t] = c[t']$$

for all observational contexts $c$. Hennicker [8, 10, 11] defines an observational predicate $\text{Obs}(t)$ for term $t \in T(\Sigma, X)$ which holds in a $\Sigma$-algebra $(A, \text{Obs}^A)$ — or $(A, \text{Obs}^A) =_{\text{Obs}} \text{Obs}(t)$ — if for all valuations $v : X \rightarrow A$ the interpretation of $t$ by $v$ is element of the observational subalgebra $(A, \text{Obs}^A)$. i.e. $v^*(t) \in \text{Obs}^A$. Models are consequently defined by:

$$\text{Mod}_{\text{Obs}}(M) := \{ A \in \text{Alg}_{\text{Obs}}(\Sigma) | A \models_{\text{Obs}} \phi \text{ for all } \phi \in E\}$$

We will present a variation of this observation predicate and observable substructures in Section 5. An implementation relation $\rightarrow$ can be defined as usually through model class inclusion, now based on observational model class construction:

$$M \rightarrow M' \text{ iff } \text{Mod}_{\text{Obs}}(M') \subseteq \text{Mod}_{\text{Obs}}(M)$$

for equal signatures of $M$ and $M'$.

Example 4.3 Define a specification $N = (\Sigma, F)$ with

$$\Sigma :$$

sorts$(\Sigma) = \{\text{state}, \text{elem}, \text{bool}\}$

attr$(\Sigma) = \{\text{top} : \rightarrow \text{elem}, \text{is\_full} : \rightarrow \text{bool}\}$

tran$(\Sigma) = \{\text{push} : \text{state} \times \text{elem} \rightarrow \text{state}, \text{pop} : \text{state} \rightarrow \text{state}\}$

$$F :$$

$[\text{push}(st,e)] \text{top}() = e$

$[\text{push}(st,e'); \text{pop}(st)] \text{top}() = e\text{true}$

Then the specification $M$ — as in Example 3.2 — is an implementation of $N$, i.e. $N \rightarrow M$. The specification $M$ is more specific since it defines the behaviour for full stacks.

We will consider implementations later on in Section 6. In order to prove correct implementations in behavioural settings, for example context induction can be applied. The induction is based on the application of observational contexts to ground terms, see work by Hennicker and Bidoit [10, 14] and also work on co-induction, e.g. [16].

5 Data Refinement

Our notion of satisfaction is already defined on the idea of observational behaviour (cf. Section 3). Internal properties of the state are invisible. They can only be inspected by attributes. The notion of observation is in particular important if implementation is taken into consideration. In implementation, concepts are realised whose properties are not relevant with respect to some abstract specification. Still, implementations shall be considered correct if they satisfy some abstract observational behaviour.

Considering the process of implementing and refining specifications and the necessity of proving the correctness of those, we find that we need means to abstract from implementation details in order to prove that implementations show
the same observational behaviour as the specifications. We have used a very loose semantics for specifications, i.e. not all carrier elements need to be representable by terms, i.e. are reachable. We have ignored reachability in order to obtain a widely defined notion of models. Non-reachable elements will be accepted, since they do not influence the behaviour of models. This allows the implementation of specifications based on libraries of existing standard implementations (for example basic data types), which might not match specifications exactly. Due to this requirement, the very loose semantics approach is sensible.

**Example 5.1** Standard models of the data type Integer can be used to implement a specification of boolean values. By applying an observational predicate \( \text{Obs}_{\{\text{true, false, not}\}}(t) \) we want to express that only the carrier elements of a carrier set Integer for sort int reachable by true, false and not are observable. These could for example be the numbers 0 and 1. All other numbers are unreachable (with boolean constructors), and therefore not observable.

The principle problem with models including non-reachable elements – that structural induction based on the syntactic structure is not possible – can be countered by using the \( \text{Obs} \) predicate.

**Example 5.2** With the formula
\[
\text{Obs}_{\{\text{true, false, not}\}}(t) \rightarrow (\text{not}(\text{not}(t)) = t)
\]
we specify that \( \text{not}(\text{not}(t)) = t \) shall only hold (and needs to be proved) in a substructure reachable with true, false and not.

We can verify in the substructures defined through the \( \text{Obs} \)-predicate. Another use of the \( \text{Obs} \)-predicate shall be introduced. \( \text{Obs} \) can be used to define constructors.

**Example 5.3** With \( \text{Obs}_{\text{int}}(\text{zero}) \) and \( \text{Obs}_{\text{int}}(n) \rightarrow \text{Obs}_{\text{int}}(\text{succ}(n)) \) we define the constructors zero and succ for natural numbers.

The term 'observational' indicates that more than reachability is aimed at. Observability allows to relax a correctness condition for implementations. The predicate \( \text{Obs}_{\Sigma'} \) can be defined for every subsignature \( \Sigma' \) of \( \Sigma \). \( \text{Obs}_{\Sigma'}(t) \) shall express that a carrier element \( v^*(\Sigma', t) \) can be denoted by a \( \Sigma' \)-ground term:
\[
\text{Obs}_{\Sigma'}(t) \text{ iff } \exists t' \in T(\Sigma'). v^*(\Sigma', t) = v^*(\Sigma', t') \text{ for } t : s \text{ and any state } ST
\]

Since we are talking about reachability or observability in the context of sorts or specific operation symbols, we offer three forms to describe the signature \( \Sigma' \) of the predicate \( \text{Obs}_{\Sigma'} \):

- \( \Sigma' : \) any subsignature \( \Sigma' \) of a given signature \( \Sigma \),
- \( \Sigma_s : \) a sort \( s \) indicating that only elements of this sort are relevant,
- \( \Sigma_{\{o_{p_1}, \ldots, o_{p_n}\}} : \) a list of selected operation symbols of a given sort \( s \).

We have seen an example of the last form – which is the most general form – in Examples 5.1 and 5.2. An example for the second form can be found in Example 5.3.

We introduce two abbreviations. Firstly, a relativised universal quantification \( \forall x : s. \text{Obs}_{\Sigma'}(x) \rightarrow \phi \) or short \( \forall x : \Sigma. \phi : \) for all elements \( \phi \) holds or they are not reachable, i.e., \( \phi \) hold for all reachable elements. Secondly, a relativised existential quantification \( \exists x : s. \text{Obs}_{\Sigma'}(x) \wedge \phi \) or short \( \exists x : \Sigma. \phi : \) there are observable element for which \( \phi \) holds. With \( \text{Obs}_{\Sigma'} \) and \( \Sigma = \Sigma' \) we model reachability.
Example 5.4 Consider Example 5.2. Instead of

\[ \forall t : s . \text{Obs} \{ \text{true, false, not} \} \rightarrow (\text{not}(\text{not}(t)) = t) \]

we would write

\[ \forall t : \{ \text{true, false, not} \} \rightarrow (\text{not}(\text{not}(t)) = t) \]

In other approaches, e.g. Hennicker’s work, contexts are used to defined observable constructs. Here, we use subsignatures.

With \( A^{\text{Obs}} \) we shall denote the observable \( \Sigma' \)-subobject for a \( \Sigma \)-object \( A \):

\[ a \in \text{Obs}_{\Sigma'}^{A} \text{ iff } \exists t \in T(\Sigma')_{a} \text{ with } a = v^\ast(ST, t) \text{ for } s \neq \text{state} \text{ and any state } ST \]

This is a standard definition based on elements reachable via ground terms of the subsignature. We construct the observable subobject as follows:

- \( A_{s}^{\text{Obs}} := \text{Obs}_{\Sigma'}^{A} \) for data sorts \( s \in (\text{sorts}(\Sigma) - \{ \text{state} \}) \),
- \( A_{\text{state}}^{\text{Obs}} := \text{Obs}_{\Sigma'_{\text{state}}}^{A} \) for sort \( \text{state} \),
- \( \text{op}_{\Sigma'}^{A} := \text{op}^\ast_{\text{Obs}} \) for any operation \( \text{op} \in \text{attr}(\Sigma) \cup \text{tran}(\Sigma) \),
- \( ST_{\text{Obs}}^{A} := ST \) for \( ST \in A_{\text{state}}^{\text{Obs}} \).

**Lemma 5.1** The observable \( \Sigma' \)-subobject of \( A \) is a \( \Sigma' \)-object.

**Proof:** Follows from the definition of a \( \Sigma \)-object and observable \( \Sigma \)-subobjects. \( \square \)

Satisfaction with respect to observable substructures can be expressed by:

\[ A^{\text{Obs}} , ST \models \text{Obs}_{\Sigma'}^{A}(t) \text{ iff } D(ST, t) \text{ for } t \in T(\Sigma') \]

**Lemma 5.2** Let \( A \) be model of a specification \( < \Sigma, E > \) and let \( t \) be a \( \Sigma \)-term of sort \( s \). Then:

\[ A , ST \models \text{Obs}_{\Sigma'}^{A}(t) \text{ iff } \Sigma' \subseteq \Sigma \text{ and } v^\ast(ST, t) \in \text{Obs}_{\Sigma'}^{A} \]

**Proof:** Follows from the definitions of the satisfaction of \( \text{Obs} \) and observable subobjects. \( \square \)

In order to illustrate how to implement a data type, we use the data type \( \text{set} \), which shall be implemented based on an existing model of tuples. The implementation shall be formalised through model class inclusion, i.e., \( M' \) implements \( M \), or \( M \hookrightarrow M' \), if \( \text{Mod}(M') \subseteq \text{Mod}(M) \). The resulting set specification will still be an abstract specification.

We will implement sets of natural numbers on tuples of integers. In the set of integers, only natural numbers shall be observable. It shall be assumed that \( \text{SET} \) and \( \text{TUPLE} \) are generic specifications which can be instantiated with element types such as \( \text{NAT} \) or \( \text{INT} \).

**Example 5.5**

\( ^{5}\text{op}^{A}_{\text{Obs}} \) is \( \text{op}^{A} \) restricted to the observable subsets of carrier sets.
spec SET(NAT) by TUPLE(INT) is
extend TUPLE(INT) by
opns
  set_insert: tuple \times int \rightarrow tuple
  set_delete: tuple \rightarrow tuple
  empty_set: \rightarrow tuple
axioms (t : tuple; i : int)
  \text{Obs}_{\{\text{empty_set, set_insert}\}_{\text{tuple}}}(t) \rightarrow 'set axioms''
  \text{Obs}_{\{\text{zero, succ}\}_{\text{int}}}(i) \rightarrow 'Peano axioms''
end spec

The operations \text{set_insert} and \text{empty_set} shall be the constructors, i.e. they denote the same reachable carriers as the whole set of operations of the signature. Only those set elements are observable which can be reached via ground terms based on \text{zero} and \text{succ}. These are the natural numbers. The axioms have to hold only for them. This is of importance if we implement on standard models of predefined types such as Integer. \text{Obs}_{\{\text{zero, succ}\}_{\text{int}}}(i) holds, if \nu^*\left(ST, i\right) is element of the subset reachable by \text{zero} and \text{succ} for any state \text{ST}. Thus, SET(NAT) by TUPLE(INT) is a correct implementation of a specification SET(NAT) with operations \text{set_insert, set_delete, empty_set} and the usual set axioms and set elements which are natural numbers that obey the Peano axioms. These are still abstract specifications, i.e., proof obligations if an executable version is derived.

The example illustrates the main application of the observability predicate: to realise a specification using another, existing one. Examples are the realisation of boolean values on a data type integer or sets on a data type list/tuple. This allows a developer to realise a specification based on suitable, existing library components. Typically, the boolean or set terms (generated by their constructors) are only interpreted in subsets of the integer or list carriers, respectively. With the \text{Obs} predicate, we can declare these subsets as observable and reason within these substructures. The technique can also be used to define export interface for existing specifications (e.g., restricting the set of constructors).

The definition of the \text{Obs} predicate has been illustrated using familiar data sort specification and implementation so far; the sort \text{state} has been excluded. The predicate \text{Obs}_{\text{state}} for sort \text{state} works in the same way as \text{Obs}_{s} for data sorts \text{s}. Applied to states it would express the reachability or observability of states, i.e., whether there is a path from an initial state to the state under consideration.

The technique of observation-based specification can also be applied to the sort \text{state}. Instead of applying the \text{Obs} predicate to constructors of data types such as boolean or set - as in the previous example - we can also use constructors of sort \text{state}. Applied to the stack object that has been specified in previous examples, these constructors would be a constant empty stack and the push and pop operations. We could realise a stack on a list object. There might be an executable implementation for a list object specification in some component library. We assume a list object specification with the usual operations (head, tail, a list constructor, concatenation) for this example.

Example 5.6

spec STACK by LIST is
extend LIST by
opns
  empty_stack: \rightarrow list
  push: list \times \text{elem} \rightarrow list
  pop: list \rightarrow list
axioms (l : list)
  \text{Obs}_{\{\text{empty_stack, push, pop}\}_{\text{list}}}(l) \rightarrow 'stack axioms''
end spec

The stack axioms are defined in previous examples. Again, the specification in the example formulates the proof obligations as axioms.
Example 5.7 More concrete specifications could contain the following axioms:

\[
\begin{align*}
\text{empty\_stack}() &= \text{empty\_list}() \\
\text{top}(st) &= \text{head}(st) \\
\text{push}(st,e) &= \text{concat}([\text{list}(e), st])
\end{align*}
\]

This describes the implementation of stacks in terms of lists more explicitly.

6 Operation Refinement

An implementation between two specifications \( M \) and \( M' \) shall be defined based on model class inclusion, i.e. we require \( \text{Mod}(M') \subseteq \text{Mod}(M) \). Model class inclusion formally captures the idea of making design decisions: there are less models if requirements are added or strengthened\(^6\). Model class inclusion guarantees that properties established for a specification are preserved in an implementation. We have not used the \( \text{Mod}_{OHS} \)-construction introduced in Section 4 here, instead we use the variant from Section 3. As we have already pointed out, this incorporates the concept of observation. How to develop a more substantial refinement approach is described in [21]. Technical details of this section including proofs can be found in [18].

An inference rule shall be introduced. The consequence rule helps to prove pre/postcondition specifications:

\[
\frac{\phi \rightarrow \phi' \land \psi \rightarrow \psi'}{\phi' \rightarrow [P] \psi'}
\]

This rule will be useful in the definition of a constructive variant of an implementation relation, called refinement, between specifications.

Example 6.1 Consider the following formula:

\[\neg \text{is\_full}() \rightarrow [\text{push}(st,e)] \text{top}() = e\]

Then the formula

\[\text{true} \rightarrow [\text{push}(st,e)] \text{top}() = e \land \neg \text{is\_empty}()\]

is a consequence of the first formula, since \( \text{top}() = e \land \neg \text{is\_empty}() \rightarrow \text{top}() = e \) and \( \neg \text{is\_full}() \rightarrow \text{true} \). The postcondition is strengthened and the precondition is weakened.

We define a few other relations between specifications. Let \( p_1, p_2 \) be transitions with the same signature and \( \phi_{p_1} \rightarrow [p_1] \psi_{p_1} \) and \( \phi_{p_2} \rightarrow [p_2] \psi_{p_2} \) be their respective specifications. The implementation \( p_1 
\rightarrow p_2 \) holds, if every function interpreting \( p_2 \) in a \( \Sigma \)-model which satisfies \( \phi_{p_2} \rightarrow [p_2] \psi_{p_2} \) is also interpretation of \( p_1 \) in a \( \Sigma \)-model which satisfies \( \phi_{p_1} \rightarrow [p_1] \psi_{p_1} \). The implementation \( M_1 \rightarrow M_2 \) holds iff \( \text{Mod}(M_2) \subseteq \text{Mod}(M_1) \) for equal signatures of \( M \) and \( M' \). The refinement relation \( p_1 \triangleright p_2 \) holds, if \( D(p_1) \land D(p_2) \) and \( \phi_{p_1} \rightarrow \phi_{p_2} \land \psi_{p_2} \rightarrow \psi_{p_1} \), hold. The consequence rule and Example 6.1 illustrate this. We will assume terminating operations for this definition such that by execution of the operations the postcondition can always be established. Let \( M_1 \) and \( M_2 \) be specifications. The refinement relation \( M_1 \triangleright M_2 \) holds, if for all \( p \in \text{tran}(\Sigma_{M_1}) \) the relation \( p_{M_1} \triangleright p_{M_2} \) holds.

Lemma 6.1 The relations \( \triangleright_{tr}, \triangleright_{sr}, \triangleright_{ts} \) and \( \triangleright_{h} \) form a partial ordering.

This is usually called the vertical composition property - a useful property for implementation relations.

In order to relate the two relations \( \triangleright_{sr} \) and \( \triangleright_{ts} \), we assume the three simplifications.

\(^6\)Certainly, it should be allowed to add new elements in the implementation which do not influence the semantics of the original specification. This can be expressed by allowing the signature of \( M \) to be a subsignature of \( M' \). In order to simplify the approach here, we have ignored this.
• There are no explicit invariants $inv^7$. They will be associated to transition definitions: reformulate $\phi \rightarrow [P] \psi$ to $\phi \rightarrow [P] \psi \land inv$.

• There are only transition specifications, i.e. box- and diamond operators and equations on sort $state$. Attributes are specified by normal first-order formulas and can be treated as invariants.

• Every transition is defined by only one formula – as justified in Lemma 6.2 below.

The following lemma needs the expressiveness of the base logic as a prerequisite – this shall be assumed here. In [18] Section 3.4, we have addressed the question of expressiveness in this context. A similar lemma can be derived for the diamond-operator.

**Lemma 6.2** Any set of formulas $\{ \phi_i \rightarrow [P] \psi_i \mid i = 1, \ldots, n \}$ can be transformed into a formula $\phi \rightarrow [P] \psi$, which holds iff all formulas $\phi_i \rightarrow [P] \psi_i$ hold.

There is a related requirement for the equational specification of hidden algebras, see [15], saying that equations on data sorts are not allowed in that framework. Any such equation has to be asserted and proved separately.

We have already seen that the refinement is compositional, i.e., the refinement between specifications depends on the refinement between the constituent transitions – by definition. The following lemma shows that the implementation is compositional in the same way. Carrier sets, the state $ST$, and functions interpreting state observers do not need to be considered.

**Lemma 6.3** Let $M_1, M_2$ be specifications with $\text{sig}(M_1) \subseteq \text{sig}(M_2)$ under the assumptions specified above. If for all transitions $t \in \text{tran}(\text{sig}(M_1))$ the inclusion $\text{mod}_{M_2}^{r \downarrow t}(t) \subseteq \text{mod}_{M_1}^{r \downarrow t}(t)$ holds, then $\text{mod}(M_2|\text{sig}(M_1)) \subseteq \text{mod}(M_1)$, or $M_1 \not\Rightarrow M_2$.

Now, we state the essential theorem of this section, i.e., that the refinement is a specialisation of the implementation.

**Theorem 6.1** Let $M_1, M_2$ be specifications under the assumptions specified above. Then it holds:

$$M_1 \models \Rightarrow M_2 \quad \Rightarrow \quad M_1 \not\Rightarrow M_2$$

The proof relies essentially on the compositionality of the underlying relations on transitions.

This theorem allows us to prove an implementation using the refinement. The refinement is a good approximation to the implementation since the only cases excluded are

• models which contain functions which do not terminate,

• formulas (describing states) that cannot be satisfied, and

• preconditions that describe states unreachable by transitions.

The refinement is only based on two non-modal first-order implications. Standard equational reasoning can be applied here.

---

7 Underlying data type and attribute specifications are considered as invariant.
7 Related Work

Our work is clearly motivated by observational and behavioural extensions to the algebraic specification approach, in particular by work from Bidoit, Broy, Hennicker, and Wirsing [7, 8, 9, 10, 11, 12, 13, 14]. Wirsing and Broy introduce ultra loose semantics, allowing behavioural abstractions by a notion of relativised quantification. This compares to the introduction of relativised quantifications $\forall x : \Sigma. \phi$ and $\exists x : \Sigma. \phi$ here in Section 5. Bidoit and Hennicker’s approach is based on a partial observational equality. Our concepts are similar to ideas of using an explicit observability predicate as presented by Hennicker in [8, 11]. We have adapted some of his ideas to the context of state-based specification and dynamic logic. Some detailed comparisons have been made in Section 5.

Other techniques of refinement and implementation for algebraically specified systems are based on a strategy of applying different implementation, restriction and other specification-building operations, see [9] Chapter 8 or [22]. We have tried to capture the idea of implementation or refinement in a single construct.

The idea of using modal logics as logics for state-based specification and implementation is not new. A few examples include the specification language COLD [23, 24] or the Modal Action Logic [25, 26] and successive work [27, 28]. Recently, modal logics have also been used in the context of coalgebras. An example is [29], where coalgebras are used to define the semantics for object-oriented programming supported by a modal logic. Coalgebras can generalise transition systems, and modal logics are a natural choice to capture this in a logic. The approach taken by the developers of hidden algebras [15, 16] – on which our semantical structures are based – follow the same direction.

We have used a dynamic logic, i.e., a modal logic that makes commands explicit. Modalities are indexed by commands. Another form of modal logics – temporal logics – are also suitable for the specification and development of state-based systems. In particular TLA (the Temporal Logic of Actions) [30, 31] offers advanced concepts for implementation and refinement. We have chosen a dynamic logic since it provides the more adequate theory for the support of the pre/post-condition technique.

Finally, we shall briefly address our own related work. The basic formal framework – excluding observational aspects – is presented in [18]. Two of our papers extend the ideas of operation refinement: [21] presents refinement ideas for the ASM (Abstract State Machines) notation, [32] is based on the idea of refinement as the matching technique for component composition.

8 Conclusions

We have presented a framework for specifying and developing state-based systems based on the idea of behavioural reasoning and observational implementation. In contrast to related approaches we have used a dynamic logic instead of an equational logic. Modal logics give us more expressive power in the specification of dynamic and reactive systems. The principle disadvantage of reasoning about implementations with the more complex modal logics is compensated using an equationally oriented refinement notion and the definition of observable substructures. One of our objectives is to develop an extension of the pre/post-condition technique for state-based systems, which would allow an improved technique to be used in component-based development approaches considering non-determinism and liveness besides the safety-specification of the classical pre- and postcondition technique. It allows us to use equational reasoning to prove dynamic systems using a refinement inference rule as the key tool.

We have seen that a behavioural notion of implementation is needed for state-based specification where we want to abstract from the realisation of the state itself. Observation is a useful tool in the realisation of specifications based on existing libraries of executable components. Our notion of satisfaction for the modal operators and equations on the state sort satisfy this requirement. Our data refinement notion is observational, based on the idea of relativisation.

We can specify dynamic state-based systems solely using modal constructors and equations of sort state. In most situations implementations can be proved using the equational refinement relation. A further step in improving the implementation of abstract specifications on predefined standard models might be using congruences to express that

14
certain elements of the implementation shall be considered equivalent, or observationally equal. Ideas from Hennicker and Bidoit might be suitable.

References


