Computer Science at Kent

Input/Output Abstraction of State Based Systems

Eerke Boiten

Technical Report No. 12-04
June 2004

Copyright © 2004 University of Kent at Canterbury
Published by the Computing Laboratory,
University of Kent, Canterbury, Kent, CT2 7NF, UK
Input/Output Abstraction
of State Based Systems

Eerke Boiten
June 2004

Abstract
Abstraction of specifications is a method of making verification and validation of specifications and implementations more tractable. This paper considers the special case where the abstraction is defined by eliding input or output variables in state based specifications — in particular, conditions for such abstractions to be sound and complete with respect to a refinement semantics. Output abstractions turn out to be unconditionally sound, and combinations of output abstractions are complete in certain circumstances. Concrete results are developed in the state-based notation Z, and then considered in the underlying semantic framework and for similar languages.

1 Introduction: Abstraction and Verification
The complexity and size of specifications and implementations form a serious impediment to verification. Model checking is able to deal with ever increasing state spaces, through advances in technology and implementation methods; however, there will always be larger or even infinite spaces of interest to be explored. Full verification, for example using refinement [7, 9], of a candidate implementation with respect to a specification, is only sparingly applied to “real” systems.

Abstraction makes these kinds of verification efforts more tractable. In model checking, where one checks whether a “model” satisfies a “property” by exhaustive search of the model’s state space, one may abstract the model. This reduces the state space to be searched, and preserves positive results: if the abstract model can be shown to satisfy a property, so does the original model. In verification, one may abstract the specification instead, leading to a preservation of negative results. Such abstractions may be viewed as
tests [1, 16], describing a subset of the originally required properties. Of course, if a candidate implementation fails the test (i.e., it produces a result inconsistent with the abstraction), then it will also fail the original specification. For appropriate definitions of refinement and abstraction (namely: set inclusion over “properties of interest”), they are each others’ converse.

This paper considers a particular kind of abstraction: namely, removing input and output variables from state based systems. The notation used is Z [15], which has a well-developed theory of refinement [9, 17].

In Section 2 we describe how Z is used as a state-based specification notation, its standard notion of refinement, and existing approaches to “abstraction as testing” in Z. Section 3 presents syntactic and semantic methods of abstracting over input and output variables. Section 4 then describes the a generalised notion of refinement, necessary to verify these abstractions, viz., IO refinement. The subsequent sections investigate “soundness”: whether abstractions are indeed converse IO refinements. For input variables, conditions for this are derived in Section 5. For output variables, an unconditional soundness result is given in Section 6. Section 7 shows that combinations of output abstractions may also be jointly “complete” in certain situations. The final section concludes, discussing how the paper’s results transfer to the underlying semantic model and other specification languages.

2 Z, Refinement, Abstraction, Testing

We first describe how abstract data types (ADTs) are specified in Z, then present the standard notion of refinement for such ADTs, and finally discuss existing approaches to abstraction and testing based on this.

2.1 Abstract Data Types in Z

State-based systems are commonly described in Z using the “states-and-operations” style. An ADT is given by a state space, an initialisation, and a collection of operations. All of these are described by Z “schemas”, which describe sets of “bindings”, essentially labelled products. The labels may be viewed as names of variables, which are indicative of their roles: primed variables represent “after-states”, inputs end in question marks, outputs in exclamation marks.

Example 1 The financial affairs of a traditional monarchy may be repre-
sented by the data type \((\text{Treasury}, \text{Init}, \{\text{Tax}, \text{Spend}\})\) where

\[
\begin{array}{l}
\text{Treasury} \\
m : \mathbb{N}
\end{array} \quad \begin{array}{l}
\text{Init} \\
\text{Treasury}' \\
m' > 0
\end{array}
\]

Initially, the treasury is non-empty. Taxing the citizens by a value \(in\) results in a corresponding increase of the treasury; the new balance is reported to the treasurer (output \(m'\)); the king only observes that the balance has increased (output \(inc\)). If there are sufficient funds, the king may request to spend an amount \(req\), which leads to the (identical) amount \(out\) being spent.

\[
\begin{array}{l}
\text{Tax} \\
\Delta \text{Treasury} \\
in? : \mathbb{Z} \\
m! : \mathbb{N} \\
inc! : \mathbb{B} \\
in? > 0 \\
m' = m + in? \\
m! = m' \\
inc! = (m' > m)
\end{array} \quad \begin{array}{l}
\text{Spend} \\
\Delta \text{Treasury} \\
req? : \mathbb{N} \\
out! : \mathbb{N} \\
out! = req? \\
m' = m - out!
\end{array}
\]

The signature \(\Sigma S\) of a schema \(S\) is defined as a schema containing all its declarations (normalised\(^1\)) with the predicate “true”. Formally, \(\Sigma S = S \lor \neg S\). Any schema \(S\) such that \(S = \Sigma S\) will be called a signature. The subsignature relation on signatures is defined by \(S \sqsubseteq T \equiv (S \land T) \subseteq T\).

Important subsignatures for an operation \(Op\) are \(?Op\) which returns the signature of the inputs, and \(!Op\) which gives the signature of the outputs. The precondition pre \(Op\) (in general not a signature) returns only the before-state and inputs, existentially quantifying over after-state and outputs.

### 2.2 Refinement

Refinement of Z ADTs is normally defined on two levels. Operation refinement, or “algorithmic refinement”, which leaves the state unchanged, is defined at the level of individual operations.

\(^1\)Normalisation replaces any declaration \(x : S\) by \(x : X\) where \(X\) is the maximal set containing \(S\), and an extra predicate \(x \in S\).
**Definition 1 (Operation refinement)** An operation $CO_p$ is an operation refinement of an operation $AO_p$ over the same state space $State$ iff

Correctness

\[
\forall State; \ State'; \ ?AO_p; \ !AO_p \bullet \text{pre } AO_p \land CO_p \Rightarrow AO_p
\]

Applicability

\[
\forall State; \ ?AO_p \bullet \text{pre } AO_p \Rightarrow \text{pre } CO_p
\]

The two conditions implicitly represent two ways in which an operation can be refined: by reduction of non-determinism, and by widening the area where the operation is guaranteed to be well-behaved, respectively. As the operations in our example are already deterministic, they can only be refined by weakening their preconditions, e.g., $Tax$ may be refined by removing the restriction $in? > 0$.

Data refinement is a generalisation of operation refinement which exploits the data type being abstract, i.e., the state may be changed provided the externally visible behaviour is preserved, and in general ADTs need to be refined in their entirety. Based on the relational refinement of He, Hoare and Sanders [11], this refinement theory for $Z$ is described in full detail in the monograph [9]. The standard method of verifying data refinement is through upward and downward simulations, which are sound and jointly complete. As the former do not contribute to this paper, we only give:

**Definition 2 (Downward simulation)** Given ADTs $A = (AState, AInit, \{AO_{pi}\}_{i \in I})$ and $C = (CState, CInit, \{CO_{pi}\}_{i \in I})$, where corresponding operations have identical input and output signatures. The relation $R$ on $AState \land CState$ is a downward simulation from $A$ to $C$ if

\[
\forall CState' \bullet CInit \Rightarrow \exists AState' \bullet AInit \land R'
\]

and for all $i \in I$:

\[
\forall AState; \ CState; \ ?AO_{pi} \bullet \text{pre } AO_{pi} \land R \Rightarrow \text{pre } CO_{pi}
\]

\[
\forall AState; \ CState; \ CState'; \ ?AO_{pi}; \ !AO_{pi} \bullet
\text{pre } AO_{pi} \land R \land CO_{pi} \Rightarrow \exists AState' \bullet R' \land AO_{pi}
\]
2.3 Refinement and Testing

Methods for test case generation based on refinement techniques often employ “horizontal” decompositions of the state space, inputs and outputs, i.e., considering partitions. For horizontal decomposition of the \textit{Spend} operation one might look at the variable $m : \mathbb{N}$ which may or may not be zero, i.e., use the disjunction

\[ m = 0 \lor m > 0 \]  \hspace{1cm} (1)

as the basis of a decomposition. In the PROST-Objects testing method described by Stepney [16], a test case might be derived by using one of these disjuncts to weaken the operation, e.g.,

\[
\begin{array}{l}
\text{SpendZero} \\
\Delta \text{Treasury} \\
\text{req?} : \mathbb{N} \\
\text{out!} : \mathbb{N} \\
\end{array}
\]

\[ m = 0 \Rightarrow (\text{out!} = \text{req?} \land m' = m - \text{out!}) \]

It is clear that \textit{SpendZero} is an abstraction of \textit{Spend}: it is only required to behave like \textit{Spend} on part of its domain.

Dick and Faivre [10] described a method of test case generation based on “disjunctive normal forms” (DNFs). Properties like (1) are used to decompose operations into disjunctions of partial operations, with each such partial operation leading to a test case. Using property (1) leads to three sub-operations, distinguishing whether $m = 0$, either before or after the operation. (The fourth case, $m = 0 \land m' > 0$, cannot arise as \textit{Spend} never increases the treasury.)

\[ \text{Spend} == \text{Spend}_1 \lor \text{Spend}_2 \lor \text{Spend}_3 \]

\[
\begin{array}{l}
\text{Spend}_1 \\
\Delta \text{Treasury} \\
\text{req?} : \mathbb{N} \\
\text{out!} : \mathbb{N} \\
\text{out!} = \text{req?} \\
\text{m'} = \text{m} - \text{out!} \\
\text{m'} > 0 \\
\end{array} \\n\]

\[
\begin{array}{l}
\text{Spend}_2 \\
\Delta \text{Treasury} \\
\text{req?} : \mathbb{N}_1 \\
\text{out!} : \mathbb{N}_1 \\
\text{out!} = \text{req?} \\
\text{m'} = \text{m} - \text{out!} \\
\text{m'} > 0 \\
\end{array} \\n\]

\[
\begin{array}{l}
\text{Spend}_3 \\
\Delta \text{Treasury} \\
\text{req?} : \mathbb{N} \\
\text{out!} : \mathbb{N} \\
\text{out!} = \text{req?} = 0 \\
\text{m'} = 0 \\
\text{m'} = \text{m} = 0 \\
\end{array} \\n\]
By construction, such sub-operations will be disjoint, and together cover the original operation. They will not, in general, be abstractions of the original operation, because they impose restrictions on after-states.

In previous work [8], we explored the interaction between data refinement and DNF-based test case generation, in particular how disjointness and covering of calculated concrete test cases may be preserved.

3 Input and Output Abstractions in Z

Complementary to the “horizontal” decomposition approaches described in the previous section, this paper explores a “vertical” decomposition, concentrating on hiding inputs and outputs only. Recall that Z ADTs are abstract, in the sense that the state variables are not directly observable. Thus, projection on IO variables rather than on state variables is more pertinent, as it refers to directly observable behaviour. In terms of testing, an abstraction over an output variable represents disregarding the value of that output; an abstraction over an input variable represents the use of an arbitrary (“randomly” generated) input value. Both of these represent simpler tests than those where all inputs should be provided and the values of all outputs should be checked, and thereby potentially a useful simplification of the testing process.

In general, removal of certain input or output variables cannot be done purely syntactically. For example, if we remove out! from Spend, with all the predicates that refer to it, we obtain

\[
\begin{array}{|c|}
\hline
\Delta \text{Treasury} \\
\text{req? : } \mathbb{N} \\
\hline
\end{array}
\]

which is not an abstraction of Spend: it removes the constraint that \( \text{req?} \leq m \). In other words, it guarantees a well-defined result in an area where Spend did not.

A more appropriate way of hiding variables is semantically based, although it can be expressed syntactically in Z, viz. through existential quantification. For example, the output out! is hidden in Spend as\(^2\):

\(^2\)In fact, an explicit hiding operator \( \setminus \) exists in Z with the same semantics.
\[ \exists \text{out} : \mathbb{N} \cdot \text{Spend} \]
\[ \Delta \text{Treasury} \]
\[ \text{req} : \mathbb{N} \]
\[ \exists \text{out} : \mathbb{N} \cdot \text{out} = \text{req} \]
\[ m' = m - \text{out}! \]

whose predicates simplify to \( m' = m - \text{req} \).

Although such “abstractions” over input and output variables are based on the standard semantics (logic and set theory), they are not therefore guaranteed to lead to converse refinements. As described in [4, 9], refinement essentially provides a second layer of semantics on top of the standard semantics. In fact, these abstractions are even guaranteed not to satisfy the refinement conditions, as they fail the condition that abstract and concrete operations have the same input and output signatures. This gap is bridged by the notion of IO refinement presented in the next section.

4 IO Refinement

IO refinement [5, 3] allows changes of inputs and outputs, and thereby changes the boundaries of the system. It is a strict generalisation of traditional Z refinement [15, 17], which does not allow such changes. Input and output form part of the observable behaviour. Thus, when performing IO-refinement we need to keep track of all changes to the inputs and outputs; this book-keeping is elided here, and for the technicalities (concerning “original input and output transformers”) we refer to [9].

Before defining IO refinement, we present the method used for modifying inputs and outputs: through composition with “IO transformers”, which are degenerate operations which have no state, just inputs and outputs.

\textbf{Definition 3 (IO transformer)} A Z schema \( S \) is an IO transformer iff \( \Sigma S = ?S \land !S \), i.e., the signature of \( S \) contains only input and output components.

For example, the schema

\[ \text{AnIT} \]
\[ \text{req?}, \text{req!} : \mathbb{Z} \]
\[ \text{req!} = \text{req?} + 1 \]
is an IO transformer: \( ?AnIT = [\text{req} : \mathbb{Z}], \!AnIT = [\text{req} : \mathbb{Z}] \).

Sometimes the converses of IO transformers need to be used; they are defined by swapping input and output roles. This is indicated by overlining, in analogy with CCS [13].

**Definition 4 (IO decorations)** For all component names \( x \), let \( \overline{x} \) be the name \( x! \), and let \( \overline{x} \) be the name \( x? \). This definition is extended to IO transformers, analogous to the normal Z schema decoration conventions. □

The converse of \( AnIT \) above is

\[
\begin{array}{c}
\text{AnIT} \\
\text{req?, req! : Z} \\
\hline
\text{req! = req?} - 1
\end{array}
\]

An IO transformer is an input transformer for an operation if its outputs exactly match the operation’s inputs, and analogously for output transformers. Particular IO transformers act as identities on the input and output side.

**Definition 5 (Input and output transformers and identities)**

An IO transformer \( T \) is an input transformer for an operation \( Op \) iff \( ?Op = \overline{!T} \) and it is an output transformer for \( Op \) iff \( !Op = \overline{?T} \).

For a schema \( S \) its input identity is defined by \( \overline{\text{Id}} S = [\overline{S}; ?S \theta?S = \theta S] \) and its output identity by \( \text{OId} S = [!S; \overline{S} \theta!S = \theta S] \). □

An input transformer \( IT \) is applied to operation \( Op \) in \( IT \gg Op \). In the absence of name capture, the meaning of this is the conjunction of \( Op \) and \( IT \), equating and hiding the matching inputs of \( Op \) and outputs of \( IT \); an output transformer \( OT \) is applied in \( Op \gg OT \).

**Example 2** The IO transformer \( AnIT \) above is an input transformer for \( Spend \), its application leads to

\[
\begin{array}{c}
\text{AnIT} \gg \text{Spend} \\
\hline
\Delta \text{Treasury} \\
\text{req? : N} \\
\text{out! : N} \\
\hline
\text{out!} = \text{req?} + 1 \\
\text{m'} = m - \text{out!}
\end{array}
\]
An output transformer for \( Tax \) is for example

\[
\begin{align*}
\text{DelInc} \\
\text{m?}, \text{m!} : \mathbb{N} \\
\text{inc?} : \mathbb{B} \\
\text{m?} = \text{m!} \\
\text{inc?} = (\text{m?} > 0)
\end{align*}
\]

and its application to \( Tax \) leads to the removal of the output \( \text{inc!} \):

\[
\begin{align*}
\text{Tax} \gg \text{DelInc} \\
\text{\Delta Treasury} \\
\text{in?} : \mathbb{Z} \\
\text{m!} : \mathbb{N} \\
\text{in?} > 0 \\
\text{m' = m + in?} \\
\text{m! = m'}
\end{align*}
\]

Derivations of conditions for IO refinement are given in [9], using the standard relational model for \( Z \). The rules derived generalise data refinement, with rules for both upward and downward simulation. In this paper we will only need the case where concrete and abstract state spaces coincide, which is covered by the accordingly restricted downward IO simulation rule below.

**Definition 6 (Downward IO simulation)** Consider ADTs \( A = (\text{State}, \text{Init}, \{\text{AOpi}_i\}_{i \in I}) \) and \( C = (\text{State}, \text{Init}, \{\text{COpi}_i\}_{i \in I}) \). Let \( IT \) be an input transformer for \( \text{COpi}_i \) which is total on \( ?\text{AOpi}_i \). Let \( OT \) be a total injective output transformer for \( \text{AOpi}_i \). \( C \) is a downward IO simulation of \( A \) iff for all \( i \in I \)

\[
\forall \text{State}; \ ?\text{COpi}_i \bullet \text{pre}(IT \gg \text{AOpi}_i) \Rightarrow \text{pre COpi}_i \\
\forall \text{State}; \ ?\text{AOpi}_i; \text{State'}; !\text{COpi}_i \bullet \\
\text{pre AOpi}_i \land (IT \gg \text{COpi}_i) \Rightarrow (AOpi_i \gg OT)
\]

**Example 3** For any suitable operation \( Op \), using \( AnIT \) above, \( AnIT \gg Op \) or \( Op \gg AnIT \) is a downward IO simulation, as \( AnIT \) represents a total bijection.
Example 4 Due to the predicate on `inc?`, `DelInc` is injective, and thus `Tax >> DelInc` is a downward IO simulation of `Tax`. Intuitively, this shows that an output variable may be removed provided its value can be derived from the other outputs. □

Particular consequences of the downward IO simulation rule are embodied in the following theorem.

Theorem 1 (Simple input and output refinement) For any ADT, adding a declaration of a new output (from a non-empty set) to one of the operations constitutes a valid downward IO simulation.

For any ADT, adding a declaration of a new input (from a non-empty set) to one of the operations constitutes a valid downward IO simulation. □

5 Soundness of Input Abstraction

In this section we investigate the circumstances in which hiding of input variables constitutes a (converse) IO refinement. We consider two data types\textsuperscript{3} with a single operation each\textsuperscript{4}:

\[
\begin{align*}
D &= (State, Init, \{DOp\}) \\
E &= (State, Init, \{EOp\})
\end{align*}
\]

where `Inp` is a signature such that `Inp \subseteq ?DOp` and the operation in `E` is obtained by abstraction over `Inp` in `D`, i.e.,

\[
EOp = \exists Inp \bullet DOp
\]

as a consequence, `?EOp`, the input signature of `EOp`, is the schema containing the remaining inputs, and we have that\textsuperscript{5} `?DOp = Inp \land ?EOp`. Another way of expressing `EOp`, using an input transformer, is

\[
EOp = (Inp \land \text{Id} (?EOp)) >> DOp
\]

In proofs, we use the fact that (partial) IO identities have no effect in piping, in particular it is also the case that

\[
EOp = \overline{Inp} >> DOp
\]

\textsuperscript{3}They are, for once, not called `A` and `C` as we will consider possible refinement in both directions.

\textsuperscript{4}There are situations, for example when considering refusals, where the restriction to a single operation allows stronger results. However, in this case the more general treatment would still be operation-by-operation, leading to a cluttered presentation.

\textsuperscript{5}Schema conjunction for disjoint signatures is really a Cartesian product.
The case where we might have expected refinement to hold is where the “concrete” specification has input variables that the abstract one does not have. However, there is a proviso.

**Theorem 2** $D$ is a downward IO simulation of $E$ if

\[ \forall \text{State}; \ ?EOp \cdot (\exists \text{Inp} \cdot \text{pre DOp}) \Rightarrow (\forall \text{Inp} \cdot \text{pre DOp}) \]

(Informally: if DOp is enabled in any state for particular input from Inp, it is enabled in that state for all inputs from Inp.)

**Proof**

The relevant input transformer is $\overline{\text{Inp}} \land \text{Id} (\exists EOp)$, which is total. The only non-trivial condition is applicability, which reduces to

\[ \text{pre} (\text{Inp} \gg \overline{\text{Inp}} \gg \text{DOp}) \Rightarrow \text{pre DOp} \]

which is equivalent to the stated condition. □

We might also state this condition as: the preconditions of the operations are independent of the values of the variables in Inp. A corresponding theorem can be proved for upward IO simulation, with the same condition.

**Example 5** Consider the Tax operation, whose precondition includes the condition $\text{in} > 0$. Thus, it does not satisfy the condition of Theorem 2, and indeed the following is not a converse refinement of Tax:

\[
\exists \text{in} : \mathbb{Z} \cdot \text{Tax} \quad \Delta \text{Treasury} \\
\text{m!} : \mathbb{N} \\
\text{inc!} : \mathbb{N}
\]

\[
\exists \text{in} : \mathbb{Z} \cdot \text{in} > 0 \\
\text{m'} = \text{m} + \text{in} \\
\text{m!} = \text{m'} \\
\text{inc!} = (\text{m'} > \text{m})
\]

We earlier observed that Tax allowed a refinement dropping the condition \text{in} > 0, and consequently also the postcondition that \text{m'} > \text{m}. The above operation does not allow such a refinement. □

For completeness’ sake, we also state the following.

**Theorem 3** $E$ and $D$ are downward IO simulation equivalent if

\[ \forall \text{State}, \text{State'}; \ ?DOp; \ !DOp \cdot (\text{pre DOp} \land \text{Inp} \gg \overline{\text{Inp}} \gg \text{DOp}) \Leftrightarrow \text{DOp} \]

□
The informal interpretation of this condition is that the input is irrelevant in \( DOp \) (note that \( \text{Inp} \gg \overline{\text{Inp}} \) represents the full relation on \( \text{Inp} \)). Obviously this implies that the input is also irrelevant in the precondition of \( DOp \).

The downward IO simulation between \( E \) and \( D \) requires the \( \Rightarrow \)-part of this condition to guarantee correctness; the reverse implication always holds.

### 6 Soundness of Output Abstraction

In this section we investigate when abstraction over output variables constitutes a converse IO refinement. We consider data types

\[
D = (\text{State}, \text{Init}, \{DOp\})
\]

\[
F = (\text{State}, \text{Init}, \{FOp\})
\]

with some signature \( Outp \subseteq !DOp \) where the operation in \( F \) is obtained by abstraction over \( Outp \) in \( D \), i.e.,

\[
FOp = \exists Outp \cdot DOp
\]

Thus, \( !FOp \) characterises the outputs of \( DOp \) that remain present in \( FOp \).

An alternative characterisation is

\[
FOp = DOp \gg \overline{Outp}
\]

Observe that the preconditions of linked operations coincide in this case:

\[
\text{pre } FOp
= \exists \text{State}' ; !FOp \cdot FOp
= \exists \text{State}' ; !FOp \cdot \exists Outp \cdot DOp
= \exists \text{State}' ; !DOp \cdot DOp
= \text{pre } DOp
\]

Whenever \( Outp \) has more than one possible value, the output transformer \( \text{OId} (!FOp) \wedge \overline{Outp} \) is not injective, and thus \( F \) can never be an IO downward (or upward) simulation of \( D \) using that output transformer. For that reason, we only need to consider whether \( D \) is an IO refinement of \( F \). Observe that Theorem 1 does not apply, as that requires \( DOp \gg \overline{Outp} \gg Outp = DOp \), which is only the case if, whenever one output is possible in any particular state, all other outputs are possible there, too.

However, there is still an unconditional positive result in this case:

**Theorem 4** \( D \) is a downward IO simulation of \( F \).
Proof
The required output transformer is \( O\text{ld} (\lnot F\text{op}) \land Outp \), taking the “abstract” outputs for \( F \) and adding \( Outp \) to those as “concrete” outputs for \( D \). This is clearly injective.

Initialisation is unaffected. Due to equality of preconditions, applicability is guaranteed. Finally, correctness requires that \( DOp \Rightarrow EO\text{p} \gg Outp \) which does indeed hold.

Example 6 In Example 4 we showed that \( \exists \text{inc}! : \mathbb{N} \cdot Tax = Tax \gg \text{DelInc} \) was an IO refinement of \( Tax \); in combination with the above result we can now conclude that they are even equivalent with respect to IO refinement.

7  Completeness of Output Abstraction

The previous sections have investigated situations where abstractions of input and output variables were valid tests of a specification with respect to IO refinement. One might view these as soundness conditions: passing such a test is a necessary condition for any implementation to be correct with respect to the specification.

In this section, we consider situations in which passing such a test is also a sufficient condition. As the input abstractions were only conditionally sound, we concentrate here on completeness of output abstractions.

Clearly, in general, we cannot expect to be able to test a system by never checking a particular output. Example 6 gave an exception: where one output is fully determined by another, we might as well not check the first output.

Viewing these abstractions as “projections” of a specification invites a geometric analogy: we might wonder what the requirements would be for the sets of variables projected onto to be a basis, i.e., when the projections together determine the specification as a whole.

We cannot characterise this directly by refinement: the specification is a refinement of each of its projections, but the reverse only holds if the specification and each projection are all equivalent. However, we can identify a specification with the set of all its refinements; if all joint refinements of all projections are also refinements of the original specification, then the projections and the original specification are “equivalent”. (Compare the construction of least common data refinements in [6].) This is characterised in the following definition.
**Definition 7 (Basis)** Consider an ADT \( D = (State, Init, \{DO_p\}) \). Let \( \{NF_j\}_{j \in J} \) be a collection of subsignatures of \( !DO_p \), and let \( F_j = \exists NF_j \cdot DO_p \) be the corresponding collection of output abstractions.

Then, \( \{NF_j\}_{j \in J} \) is a basis of \( D \) iff for all ADTs \( H = (State, Init, \{HO_p\}) \) with input and output signatures identical to \( D \), whenever \( H \) is a downward IO simulation of each of \( F_j \) with respect to output transformer \( NF_j \), \( HO_p \) is an operation refinement of \( DO_p \).

Because \( NF_j \) is an output signature, the operation \( F_j \gg NF_j \) has the same behaviour as \( DO_p \), except for producing arbitrary values for outputs in \( NF_j \). Thus, \( F_j \) and \( F_j \gg NF_j \) are equivalent when interpreted as predicates in a context where \( \Sigma DO_p \) is defined.

Note that a basis is represented by sets of variables that are hidden rather than their complements, the variables projected onto. In particular, including the empty subsignature (i.e., projecting onto the full set of variables) will always lead to a basis.

Using the particular relation between \( DO_p \) and \( F_j \), we can eliminate the quantification over all \( H \) in Definition 7:

**Theorem 5 (Basis condition)** For nonempty \( J \), \( \{NF_j\}_{j \in J} \) is a basis for \( D \) if

\[
\forall State; State'; ?DO_p; !DO_p \cdot (\forall j : J \cdot \exists NF_j \cdot DO_p) \Rightarrow DO_p
\]

**Proof** We need to prove that \( HO_p \) is an operation refinement of \( DO_p \), using the fact that \( H \) is a downward IO simulation of each of \( F_j \). Recall that the conditions of operation refinement are

\[
\forall State; ?DO_p \cdot pre DO_p \Rightarrow pre HO_p
\]

\[
\forall State; State'; ?DO_p; !DO_p \cdot HO_p \land \pre DO_p \Rightarrow DO_p
\]

The first (“applicability”) condition requires that \( J \) is non-empty. From IO-refinement from \( F_j \) to \( H \) (with the identity input transformer), we get

\[
\pre F_j \Rightarrow \pre HO_p
\]

and we also have that

\[
\pre F_j = \pre DO_p
\]

by the construction of \( F_j \). Together these prove the first condition.
For the correctness condition, we have from IO-refinement that

∀ j : J • \( \text{pre } F_j \land HOp \Rightarrow F_j \gg NF_j \)

which, using \( \text{pre } F_j = \text{pre } DOp \) is equivalent to

∀ j : J • \( \text{pre } DOp \land HOp \Rightarrow F_j \gg NF_j \)

and thus to

\( \text{pre } DOp \land HOp \Rightarrow \forall j : J \cdot F_j \gg NF_j \)

Assuming the condition of the theorem, this proves the correctness condition.

The interpretation of the basis condition is: outputs can be verified independently, provided that all constraints on their values can be defined in terms of the values of state variables and inputs only.

**Example 7** As we already showed that Tax and Tax \( \gg DelInc \) are equivalent, it follows that projections on inc! and m! separately or jointly, or on m! only, are bases for Tax.

Note that the values of these two outputs are *not independent* of each other; however, in any given state, the value of each can be determined independently from the one from the state only.

**Example 8** For an operation which has no output projections which form a basis apart from itself, consider the operation where the king’s treasury is non-deterministically split between his sons.

$$
\begin{array}{c}
\text{Succession} \\
\Delta \text{Treasury} \\
\text{william!}, \text{harry!} : \mathbb{N} \\
\hline \\
\text{william!} + \text{harry!} = m \\
m' = 0 \\
\end{array}
$$

The projection on one of the outputs (the other is symmetric) is:

$$
\exists \text{william!} : \mathbb{N} \cdot \text{Succession} \\
\Delta \text{Treasury} \\
\text{harry!} : \mathbb{N} \\
\hline \\
\text{harry!} \leq m \\
m' = 0
$$
and the conjunction of the two projections is weaker than Succession: it contains \( \text{harry!} \leq m \land \text{william!} \leq m \) rather than \( \text{harry! + william!} = m \). Thus, the collection of projections on individual outputs fails the basis condition. 

\[\square\]

8 Concluding Comments

The results in this paper were developed in \( Z \), using a non-standard refinement relation. The outcomes were reasonably intuitive: hiding an output is always sound, hiding an input may not be; outputs can sometimes be verified independently. In the underlying semantic framework \([9, 11, 17]\), inputs and outputs are included in sequences, which are both part of the “hidden” local state and the “visible” global state. As such, they play a similar role to “normal” \([7]\) or visible variables in guarded command language based refinement, and our results should carry over to such a context. In the context of (deterministic) programming languages, projections over such sets of visible variables are a special case of slicing \([14]\), using a statement that exhibits their values. Slicing is an accepted basis for program verification and testing in that context \([2, 12]\).

Finally, note that most of our results were developed for data types with a single operation. This is no fundamental restriction, as the simple kinds of IO refinement considered lead to conditions per individual operation. Analogous results for ADTs with multiple operations would be obscured by the abundance of \( i \)-indexed sets of projections and IO transformers.

Acknowledgement

John Derrick provided useful comments on an earlier version of this paper.

References


**Note**

Section 7 can probably be simplified by first proving that every \( HOp \) in Definition 7 is an operation refinement of \( \forall j : J \bullet F_j \gg NF_j \), which I suspect is the least common IO refinement of all \( F_j \). Note that the reverse implication of the basis condition is universally true. All in all this should lead to "iff" in Theorem 5.