



Developing Components for the Common Modelling Protocol

Version: 9 September 2005

Developing Components for the Common Modelling Protocol

Contents

1. Introduction	2
2. Simulations, models and sub-models	3
3. A short introduction to the Common Modelling Protocol	5
3.1. Components and modules	5
3.2. Properties	5
3.3. Events and event handlers	5
3.4. Systems and simulations	6
3.5. Messages	6
3.6. Identifiers for entities in a simulation	6
3.7. Defining data types and units: DDML	7
3.8. Simulation documents: SDML	7
4. Designing a component: step by step	8
4.1. Purpose and boundaries	8
4.2. Identifying the submodel interface - properties	9
4.3. Identifying the component interface - events	10
4.4. Component description documents	10
5. Implementation basics	13
5.1. Required binaries	13
5.2. Component-building API: the <i>TFarmWiseInstance</i> class	13
5.3. Event processing using the state machine	15
6. Implementing a component: step by step	17
6.1. Defining the properties	17
6.2. Defining the event handlers	18
6.3. Initialising properties	18
6.4. Acquiring values for driving properties	19
6.5. Implementing model logic – event handlers	19
6.6. Providing values for owned properties	20
6.7. Publishing events	21
6.8. Handling errors	21
6.9. Discovering the current time step	22
7. Advanced topics	23
7.1. The connection scheme	23
7.2. Querying the simulation for its structure	23
7.3. Other <i>TFarmwiseInstance</i> methods that generate messages	23
8. Further documentation	24
8.1. Common Modelling Protocol specification	24
8.2. Typed Value specification	24
8.3. Help file	24

1. Introduction

This document is designed to assist builders of components that execute within the CSIRO Common Modelling Protocol (CMP). Topic areas covered are:

- an introduction to simulations;
- an introduction to CMP components;
- a description of the process by which components should be designed and implemented; and
- information about the infrastructure code provided by CPI to implement components.

Design of a component is properly the role of a “scientist”; the most important design questions relate to the way in which the various quantities and equations in a model should be structured. Implementation of a component, on the other hand, is the role of a “programmer”. Even if the same person carries out both design and implementation, it is useful to distinguish the two roles.

This document is based on the CMP implementation written by CSIRO Plant Industry (CPI). While the software development processes and examples described here are specific to the CPI implementation of the protocol, the “science” part of the component-building process should apply regardless of which implementation of the CMP is being used.

2. Simulations, models and sub-models

The purpose of any CMP component is to implement a *sub-model* that can be coupled with other sub-models to form a dynamic model. In order to develop components efficiently, a component designer must understand the concepts of *model*, *quantity*, *event* and *simulation*.

- A *simulation* is a computation of a *dynamic model* between given start and end times, i.e. it is an integration over time.
- A dynamic model is defined by a set of *equations*. The equations of a dynamic model may fall into natural groupings known as *sub-models*. Some of these sub-models may have equations and quantities of identical form, i.e. they belong to the same *sub-model class*. The dynamic model as a whole can therefore be viewed as a collection of instances of various sub-model classes.
- A sub-model is composed of a set of *quantities*, a set of *rate equations*, and a set of *events*.
- All quantities can be expressed as real numbers, integers, or Boolean values. (In the modelling protocol, these quantities can be organized into arrays and structures.)

Real-valued quantities have *dimension* and *units*; the units must conform to the dimension. When designing a sub-model, it is useful to distinguish the following kinds of quantities:

- (a) *Constants* are quantities that are (i) invariant in time and (ii) have the same value in all instances of a sub-model within a model and all simulations of a model. For example, Avogadro's number is a constant.
 - (b) *Parameters* are quantities that are invariant through the time course of a simulation, but may take different values between different instances of a sub-model within a model or between simulations of a model. For example, the radiation use efficiency of a plant species under reference conditions is treated as a parameter in many plant growth models.
 - (c) *State variables* are quantities that may vary in time as the simulation is computed. The value of a state variable must be stored in order to compute the dynamics of the sub-model. An example is the water content of the surface 100mm of soil in a moisture budget. The initial value of each state variable must be specified in order for the simulation to be computed. There is a one-to-one correspondence between state variables and the rate equations of the sub-model. In principle, there should be no redundancy in the state variables.
 - (d) *Summary variables* may also vary in time, but their value at any given time may be determined from the current values of the state and driving variables. Summary variables may be used to provide output from the simulation; to provide driving variables for other sub-models; or as notational conveniences in the specification of the sub-model's rate equations (in which case they may be referred to as "intermediate" variables). For example, in a model where the masses of different plant parts are followed the total shoot mass of the population is a summary variable.
 - (e) *Driving variables* are quantities which are stored externally to a given sub-model but which must be known in order to compute the dynamics of the sub-model. They may (and usually do) vary in time. For example, minimum daily temperature and maximum daily temperature may be driving variables of a plant growth model.
- Each real-valued state variable has a *rate equation* associated with it. The rate equation is an ordinary differential equation that gives the rate of change of that state variable over time. The right-hand side of each rate equation must be composed only of constants, parameters, state variables, summary variables and driving variables belonging to the sub-model.
 - Each submodel has zero or more *events*¹ associated with it. An event, in this sense, is
 - a set of equations defining an instantaneous change in one or more state variables; and

¹ N.B. The term "event" is used in another sense within the modelling protocol. Model events will be coded as protocol event handlers, but so will other model calculations. Unfortunately no good alternative term exists.

- a "trigger": a logical relation that, if satisfied at any time, causes the change(s) in state variables. Each event has zero or more quantities, known as *event parameters*, that may be used in specifying the right hand sides of the equations and the trigger along with constants, parameters, state variables, summary variables and driving variables belonging to the sub-model.

- A simulation is therefore completely defined by:
 - the model, i.e. the set of sub-models it contains;
 - the start and end times for the computation;
 - the values of the state variables and event parameters of each sub-model at the start time;
 - the time course of the model's driving variables.

3. A short introduction to the Common Modelling Protocol

3.1. Components and modules

Simulations that are computed within the CMP are constructed from *components*. Most components embody a sub-model.² An instance of a component within a simulation is known as a *module*. There may be many modules of the same component within a single simulation; for example, in a simulation with several paddocks there may be a water balance module for each paddock. Each module has a case-insensitive *name* (e.g. `water_b` in Figure 3.1).

3.2. Properties

Properties correspond to the quantities in a dynamic model. Every property forms part of a component. Each property is defined by a name; a type; and a value. Properties are referred to by their name: where necessary, the property name is qualified by the name of the module to which it belongs.

The type of a property determines the set of values it may take and the units of those values, where applicable. The type must be either one of a set of *primitive types* (usually real, integer, Boolean or text) or else an array or record structure ultimately composed of these primitive data types. The value must conform to the type. Type information is conveyed using an XML notation called Data Description Markup Language (DDML), which is described in section 3.7.

For example, in the simulation structure shown in Figure 3.1, `paddock_1.area` might have units of ha and a value of 150.0, while `paddock_2.area` would have the same units but a value of 350.0.

Two different kinds of properties are distinguished. *Driving properties* correspond to driving variables. The systems in a simulation work together to ensure that the values of each module's driving properties are read from the properties of other components as needed. A request for a driving property may result in zero, one, or more than one values being returned from different components, depending on the structure of the simulation.

All other properties (i.e. those that are stored by the component) are known as *owned properties*.

Some of the owned properties (usually those corresponding to state variables) have to be provided in order to define the starting state of a simulation: these are called *initialisation properties*.

3.3. Events and event handlers

Events in the common modelling protocol are used to signal the occurrence of activities (i.e. computations) and to pass instructions between components. Protocol “events” are therefore used to drive the computation of both continuous processes and logical “events” (discontinuous processes). Every event has a name; it may also have a type and parameter data that conform to the type. When an

Examples of valid units for real variables:

g	base (S.I.) unit
hPa	scaled unit
MJ/m ² /d	ratio with two terms in the denominator
/s	no numerator
kg ^{0.75}	or kg ^{.75}
m ^{1/3}	but not m ^{1/2} , which is grammatically correct but should be given as m ^{0.5}
g.m/s ²	but not m/s ² .g as numerator terms must precede denominator terms

For more detail about how to specify units, see section 6.2 of the full CMP specification.

² Other components carry out utility tasks such as controlling the order of computations within each time step, the acquisition of the values of driving variables (for example weather data) or the storage of simulation results for later analysis. This document is primarily intended for developers of components that implement sub-models.

event is passed to a component, an *event handler* is executed.

3.4. Systems and simulations

A *system* is a component that groups related components within a simulation. In addition to the usual attributes of components, a system has zero or more components within it. The components in a simulation are therefore organized into a tree structure. At the root of the tree is the *simulation system*.

3.5. Messages

Messages are the means by which information and requests are passed between components and systems as a simulation is computed. Message data are transmitted in a binary format. There is a defined set of 31 messages. Every message has a common header that contains the message type (as a numeric index), the component from which it originated and the component to which it is to be sent. Most of the messages also contain further data, which are laid out in a specified fashion so that any protocol-compliant component can read them. A component that receives a message may execute some of its own internal logic (which may result in the component sending further messages); or it may be required to send particular messages as a mandatory response.

Component builders will not generally have to work with messages directly. A component API (implemented as a class) takes care of decoding incoming messages and encoding outgoing ones.

3.6. Identifiers for entities in a simulation

Any component, property or event can be uniquely identified within the simulation by its *fully-qualified name*. The fully-qualified name is constructed by adding the names of all containing systems and (where appropriate) components before the name of the entity, delimited with “.”. For example, in Figure 3.1 the water balance component may have a property called *sw*; the fully-qualified name of this property in the *water_a* module would be *paddock_z1.water_a.sw*.

In addition to their names, all modules, properties and events within a simulation system have unique numeric identifiers. Each module is allocated a unique integer identifier during the initialisation of the simulation. Each property within a module has its own integer identifier, which is usually assigned by the component builder. The ordered pair [component ID,property ID] therefore uniquely identifies a property within the simulation. A similar scheme applies to event handlers.

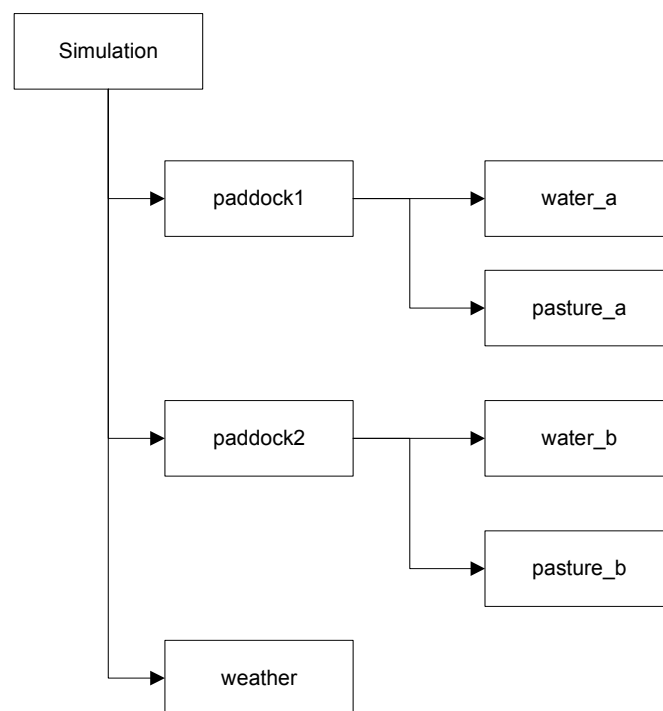


Figure 3.1. Components and systems in a CMP-compliant simulation.

3.7. Defining data types and units: DDML

Every property and event parameter has a type and (if a numeric real value) a unit. Within the protocol, information about types and units is given using Data Description Markup Language (DDML). DDML is an XML-based format; it is described in section 6 of the protocol specification document.

No knowledge of DDML is required to create properties that are scalars or arrays of scalars; the component-building API generates the DDML automatically for these types. For properties that contain multi-dimensional arrays or records, well-formed DDML descriptions of the type are required.

Examples of DDML definitions:

```
<type kind="double" unit="mm"/>           Double-precision scalar – note that the unit is given
<type kind="string" array="T"/>         Array of text strings
<type>                                   Record containing an integer, a real value and a logical value
  <field name="field_1" kind="integer4"/>
  <field kind="field_2" kind="double" unit="-"/>   Dimensionless value
  <field kind="field_3" kind="boolean"/>
</type>
<type array="T">                         Array of records...
  <element>                                ... each record has a string and two arrays of numbers
    <field name="crop_ident"                kind="string"/>
    <field name="layers"                    unit="mm" kind="double" array="T"/>
    <field name="uptake"                    unit="mm" kind="double" array="T"/>
  </element>'
</type>
```

3.8. Simulation documents: SDML

Simulation Description Markup Language (SDML) is used within the protocol to specify the structure of a simulation and the initial values of the properties in each component. A valid SDML document is a complete specification of a single simulation. SDML is an XML-based format; it is described in section 7 of the protocol specification document.

Component developers will not generally need to handle SDML directly. The protocol implementation parses the simulation structure and creates the necessary modules automatically. The component-development API parses the initialisation data given in the SDML document and provides it to the component's code as *TTypedValue* objects (see section 5.3).

4. Designing a component: step by step

Adding a new model of biological or physical processes to the protocol typically follows three steps:

1. **Define the purpose & hence the “boundaries” of the sub-model.** The component designer needs to answer the following questions:
 - What quantities should the sub-model contain?
 - What quantities belong in other sub-models?
 - What calculations is the sub-model responsible for?
2. **Specify the interface of the sub-model.** Once a conceptual boundary has been drawn around a sub-model, it becomes possible to identify the flows of information that must take place as the sub-model is set up and its equations are computed. This information specifies the interface to the component. **We strongly recommend that the interface be formally specified** in a “component description document”.
3. **Implement the sub-model as a component.** This process is described in sections 5 and 6.

These three steps perform a translation from a sub-model (a mathematical entity) to a component (a software entity). The roles required shift accordingly:

- Step 1 should be carried out by the “scientist” (the person responsible for the model equations).
- Step 2 should be carried out jointly by the “scientist” and the “programmer”.
- Step 3 should be carried out by the “programmer”.

In practice, of course, the component development process tends to be iterative; a component’s interface evolves along with the scientific understanding that its sub-model embodies.

4.1. Purpose and boundaries

This step involves making an assessment of how the new processes relate to one another and to existing processes that are already modelled within other protocol-compliant components.

When a new sub-model is being added to a suite of existing components, its boundaries will generally be evident. Where an existing model is being re-implemented within the CMP, however, the component designer must decide whether to implement the model logic as one or as several components and where to set the boundaries between the new components.

The following points should be taken into account when delimiting component boundaries:

- The most important consideration is that equations that are tightly coupled (in the sense of sharing common quantities) should form part of the same component.
- Where a component is intended to be used in conjunction with other components, their conceptual boundaries should be aligned: no process should be represented in both components and all necessary processes must be represented in one of them.
- The component designer should aim to align component boundaries with those of existing components that perform the same or a similar function. This will facilitate component interchange.

Delimiting component boundaries: special cases

- Sometimes the question of the sub-model that should represent a given process is a matter of scientific debate, and a perfect alignment of component boundaries is therefore not feasible. (This usually happens at physical interfaces such as that between soil and roots.) In this case we suggest that a new component should be implemented so as to detect its companion components at run-time and to adjust its computations accordingly.
- Components with similar processes can have boundaries that align without being identical; two or more components may together carry out the functionality of a single alternative component. For example, in crop modelling the dynamics of dead surface residues are typically considered separately from that of green biomass, and the two sets of processes are implemented as separate components. In grasslands, however, both green and dead herbage are sources of animal feed and so it makes more sense to include them in a single component. To preserve alignment of the boundaries, the interface to the grassland component needs to correspond to that of both a crop and a surface residue component (except for the elements that communicate between the two).

- There is a design tradeoff to be made between having many small components, which will increase flexibility, and having few large components, which will increase computational efficiency. Also, designing components that are too small will make it difficult to configure simulations correctly; designing components that are too large will result in users having to define unnecessary initial values.

4.2. Identifying the submodel interface - properties

Once the conceptual boundary has been drawn around a sub-model, it becomes possible to identify the full set of quantities in the model equations. The different kinds of quantities will be handled differently in the component code:

Constants	Constants will generally be defined inside the model logic code and will not appear in the component's interface.
Parameters	Values for parameters will need to be provided at startup. The component will require either an <i>initialisation property</i> for each parameter, or else initialisation properties that allow the component to locate and read file(s) containing some or all of the parameters.
State variables	Initial values for state variables will need to be provided at startup. It is good practice to define an initialisation property for each state variable, so that the initial values for a simulation are recorded in the simulation's SDML document (see section 3.8). It is also good practice to make each state variable <i>readable</i> ; code will need to be written to provide its current value to the rest of the simulation upon request.
Summary variables	Summary variables have different uses: <ul style="list-style-type: none"> • as outputs; • as driving properties by other components. In either case summary variables must be made readable if they are to carry out their function.
Driving variables	The <i>driving property</i> corresponding to each driving variable must be given a name and type that are the same as (or at least compatible with) those used in the component(s) that will provide values of the driving variable when simulations are run.

It is good practice to ensure that all initialisation properties are also readable; this simplifies the process for creating and restoring checkpoints.

The following points should be taken into account when defining the names and types of properties:

- One of the major benefits of the CMP is that it allows interchange of different models for the same or similar processes. This benefit will only be realised if properties with the same meaning are given the same definition. Component interface designers should therefore consult the documentation for related components and use common definitions wherever possible.
- Summary properties that are intended for use as outputs should be given a type and units that are convenient for users.
- Summary properties that are to be used as drivers for other components must be given a name and type that are the same as those used in the receiving component. (The CPI protocol implementation will carry out some type and unit conversions; see sections 8.4 and 6.2.3 of the full CMP specification for details.)

Writeable properties

Properties may be defined as being *writeable* as well as readable. Another component can request that a writeable property be reset to a different value.

Property resets should be used sparingly, if at all. As a rule, the only context in which they are defensible is when the period of scientific interest is after the start of the simulation and the resets are used to shift part of the system state to a desired point. Use of property resets in the regular computations of a component is a sign that its boundaries or events are badly defined.

4.3. Identifying the component interface - events

In the CMP, all computations within each time step are carried out in response to an event. When defining the component interface, it is therefore necessary to work out which equations should be computed together and so to group the model logic into *event handlers*.

Conceptually speaking, event handlers can be divided into three kinds:

- those which implement one or more rate equations and which should therefore be computed exactly once in each time step;
- those which implement the equations of an “event” in the sense of section 2, which may be computed zero or more times in each time step; and
- those that transmit information (for example the amount of a flow of organic matter).

The first kind are referred to as *sequenced event handlers*, because in the CPI protocol implementation the events that trigger them are usually generated by a special “sequencer” component that controls the computation of each time step. We recommend that sequenced event handlers be considered separately when documenting a component’s interface.

The view within CPI is that information transmission should be carried out by means of property values, not events. The distinction between using an event to trigger a submodel “event” and using an event to transmit information is not a sharp one, however; even when an event is intended to trigger a computation, its parameters convey information to the component that receives it.

While the three kinds of event handlers are logically distinct, they are implemented in the same way within the component code.

The following points should be taken into account when defining the names and types of event handlers:

- As with properties, maintaining consistency between components (both those with similar event handlers and those that generate events with similar purpose) is an important design criterion.
- Where an event represents a management intervention, its parameters should be defined in terms that are meaningful to the user.

Computation order within time steps

Some components will need information that is computed by other components during the same time step. One example is the rate of some related process; another is the case where the source component reads time-varying information in from an external file. When this situation arises, it is important that computations take place only once the correct values of the necessary driving variables have been calculated.

In the CPI protocol implementation, this is generally done by specifying a “sequencing order” for each of the events that is triggered by the sequencer component

4.4. Component description documents

CPI recommends that the process of specifying the interface to a component should be formally documented in a fashion similar to that shown overleaf.

A sample component description document

This sample describes a real (if small) component:

1. Purpose of Component

The SOILT component encapsulates the EPIC soil temperature model (Williams 1985), as modified by Potter & Williams (1994).

2. Initialisation Properties

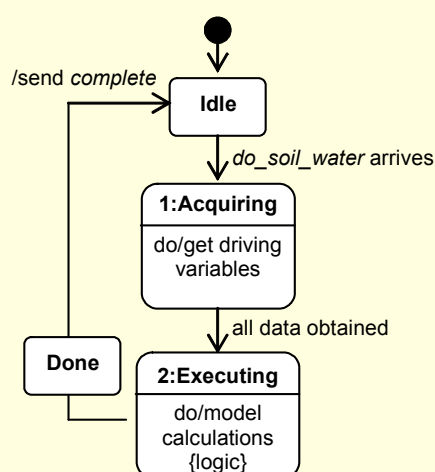
Property	Type	Units	Required?	Description
<i>ann_temp</i>	double	°C	Yes	Long-term average annual temperature at the location of the simulation.
<i>layers</i>	double[]	mm	No	Thickness of each soil layer. If not provided, layer depths will be taken from the <i>bd_layer</i> driving property.
<i>soilt</i>	double[]	°C	No	Initial soil temperature in each soil layer. If not provided, all layers will be initialised to the value of <i>ann_temp</i> .

3. Subscribed Events – Sequenced

3.1. *do_soil_water*

Default sequencing: 3000

Computes the soil temperature model for the time step.



4. Subscribed Events – Other

None.

5. Published Events

None.

6. Driving properties

Property	Type	Units	Event:State	Number*	Description
<i>bd_layer</i>	record		initialisation	1	Soil bulk density profile.
: <i>layers</i>	double[]	mm			
: <i>bd</i>	double[]	Mg/m ³			
<i>cover_tot</i>	double	m ² /m ²	<i>do_soil_temp</i> :1	0+	Total cover of co-occurring plant species.
<i>daylength</i>	double	hr	<i>do_soil_temp</i> :1	1	Day length including civil twilight.
<i>maxt</i>	double	°C	<i>do_soil_temp</i> :1	1	Maximum air temperature.
<i>mint</i>	double	°C	<i>do_soil_temp</i> :1	1	Minimum air temperature.
<i>residue_cover</i>	double	-	<i>do_soil_temp</i> :1	0+	Cover of standing dead and litter.
<i>snow_pack</i>	double	mm	<i>do_soil_temp</i> :1	0-1	Snow on the soil surface, in water equivalents. Default is 0.0.
<i>sw_layer</i>	record		<i>do_soil_temp</i> :1	1	
: <i>layers</i>	double[]	mm			Thickness of each soil layer.
: <i>value</i>	double[]	mm/mm			Current soil water content in each soil layer.
<i>shoot_dm</i>	double		<i>do_soil_temp</i> :1	0+	Total dry weight of all herbage.

* "Number" refers to the number of sources for the driving variable that is permitted.

A sample component description document (continued)

7. Owned properties

All initialisation properties are readable. In addition, the following owned properties are available:

(a) Standard properties

Property	Type	Units	Description
<i>name</i>	string		Fully-qualified name of the component.
<i>type</i>	string		Value is “Soil Temp”
<i>version</i>	string		Value is “1.0”
<i>author</i>	string		Value is “CSIRO Plant Industry”
<i>active</i>	Boolean		Denotes whether or not the component is active
<i>state</i>	string		SDML description of the current state

(b) Component-specific properties

Property	Type	Units	Description
n/a			

Constructing and distributing such “component description documents” has significant benefits.

- It enables the “scientist” to provide the “programmer” with a clear account of the code that must be written in order to implement the component wrapper (see section 6).
- The descriptions of existing components are important resources for the developers of new components, because they provide the information about the kinds and types of properties and events that enable consistency of definitions to be achieved.

5. Implementation basics

In the CPI implementation of the protocol, each component is implemented as a Microsoft Windows DLL. The implementation of a component interface as computer code should be a routine task once the sub-model's interface is properly specified and documented and its equations are correctly coded.

5.1. Required binaries

Running simulations using the CMP requires a number of executable files in addition to the CMP-compliant components. The CPI protocol implementation is structured as shown in Figure 5.1.

The central part of the CPI protocol implementation is `prot.dll`. This DLL contains the *protocol engine*, which manages the messages that communicate between the modules in a simulation. The User Interface is built to allow the user to customise a simulation and display results. `piwrapper.dll` provides an interface that allows the protocol engine to communicate with component DLLs that have been constructed with the *TFarmwiseInstance* API (see below). The `sim.dll`, `timesvr.dll` and `sequencer.dll` components are supplied by CPI. They each carry out necessary tasks in the computation of a simulation, but they could be replaced by a component developer who wished to carry out their tasks in a different fashion. `ddlodb.dll` is an optional DLL that allows a single access point to a database of structured type definitions.

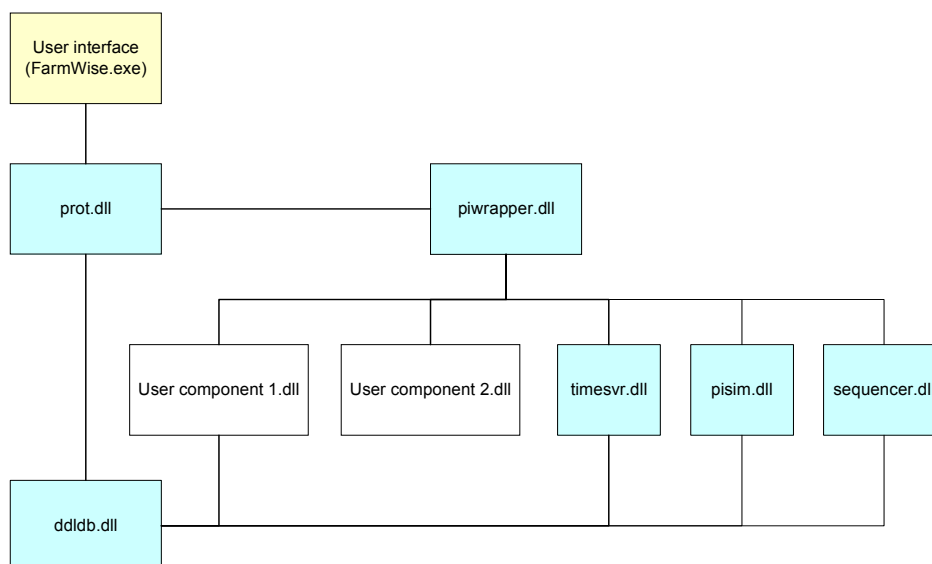


Figure 5.1. Structure of the CPI protocol implementation: executable files.

TFarmwiseInstance has been implemented in both C++ and in Delphi (Object Pascal). Examples in this document are given in Object Pascal. The class is distributed as compiled code (.OBJ for C++ and .DCU for Delphi). Details of the interface to the class (including header files for C++) are provided along with the compiled code.

5.2. Component-building API: the *TFarmwiseInstance* class

An efficient way to build CMP-compliant components, especially if they are to be executed within the CPI implementation, is to use CPI's component-building API. This API is provided as compiled code for a set of classes that implement most of the CMP-specific functionality. By using these classes, a component builder can concentrate on coding the sub-model-specific logic.

It is good practice to keep the interface code and the code containing model logic separate. As a rule, a "logic class" containing the model equations will be written either prior to, or in parallel with, the implementation process described in section 6.

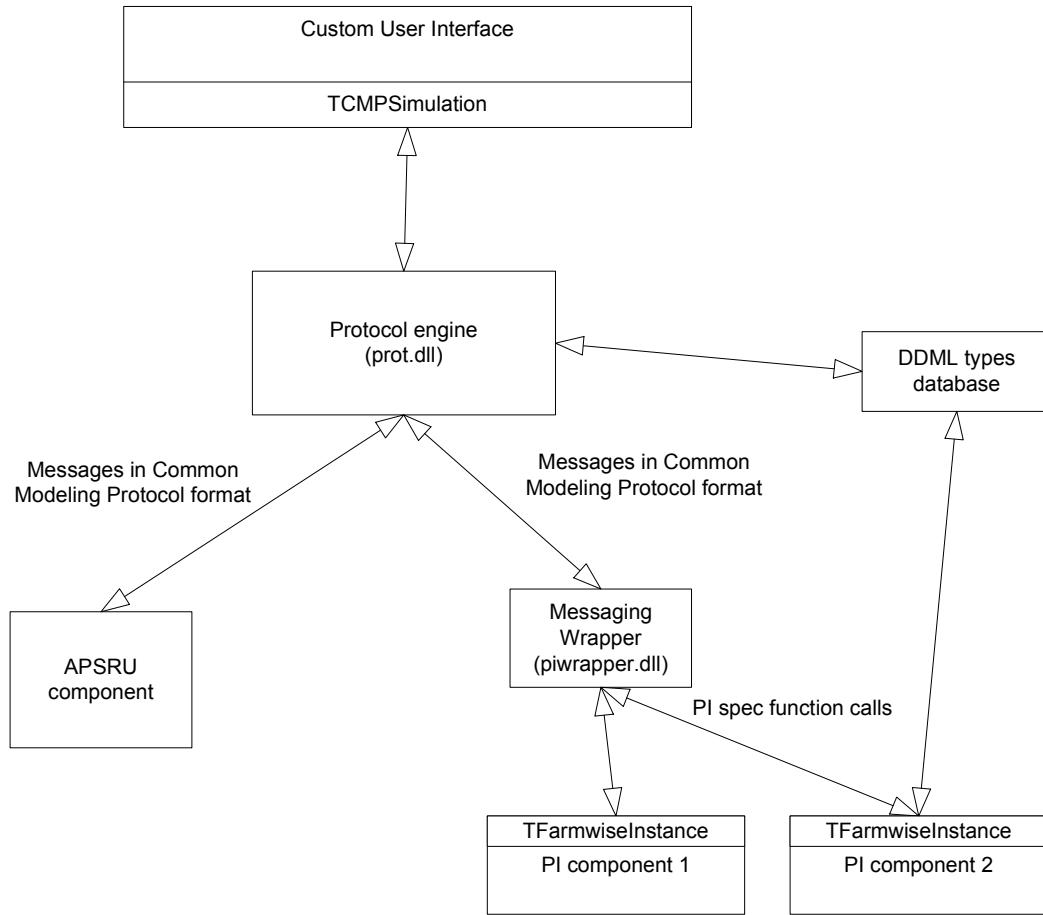


Figure 5.2. Structure of the CPI protocol implementation: view of the key classes and communication pathways. The process for linking the protocol engine to a user interface via the *TCMPSimulation* API class is described in a separate document.

The supporting classes are available in C++ and Pascal. Development tools currently supported include Visual C++, Borland C++ Builder and Borland Delphi.

The main parent class in the component-building API is called *TFarmwiseInstance*. This class decodes and encodes messages, initialisation information etc and provides the information they contain in readily-accessible forms. Almost all the work in implementing a component is done by creating a descendant class of *TFarmwiseInstance*.

Figure 5.3 illustrates the internal structure of two components implemented with *TFarmwiseInstance*. On the left, a *TTextReporter* logic class has been “wrapped” inside a *TFarmwiseInstance* descendant called *TTextOutInstance* that provides the interface to the rest of the protocol implementation. *TTextOutInstance* contains code to handle definition of the component properties and event handlers, initialisation, the acquisition of driving variables and provision of owned variables and the handling of events. On the right, a second logic class (*TLocalityModel*) has been “wrapped” inside a second *TFarmwiseInstance* descendant called *TWeatherInstance* that performs corresponding tasks.

The *TTypedValue* class hierarchy is used to provide data values (including various types and units). *TTypedValue* itself is an abstract class; it is implemented as *TSDMLValue*, *TDDMLValue* or *TInitValue*.

Because the CMP depends upon XML, parsing support for XML is also required. The C++ version of the component-building API requires Xerces 2.5.0, while the Delphi version requires XDOM 2.3. Both these XML parsing systems are open source; they can be located at

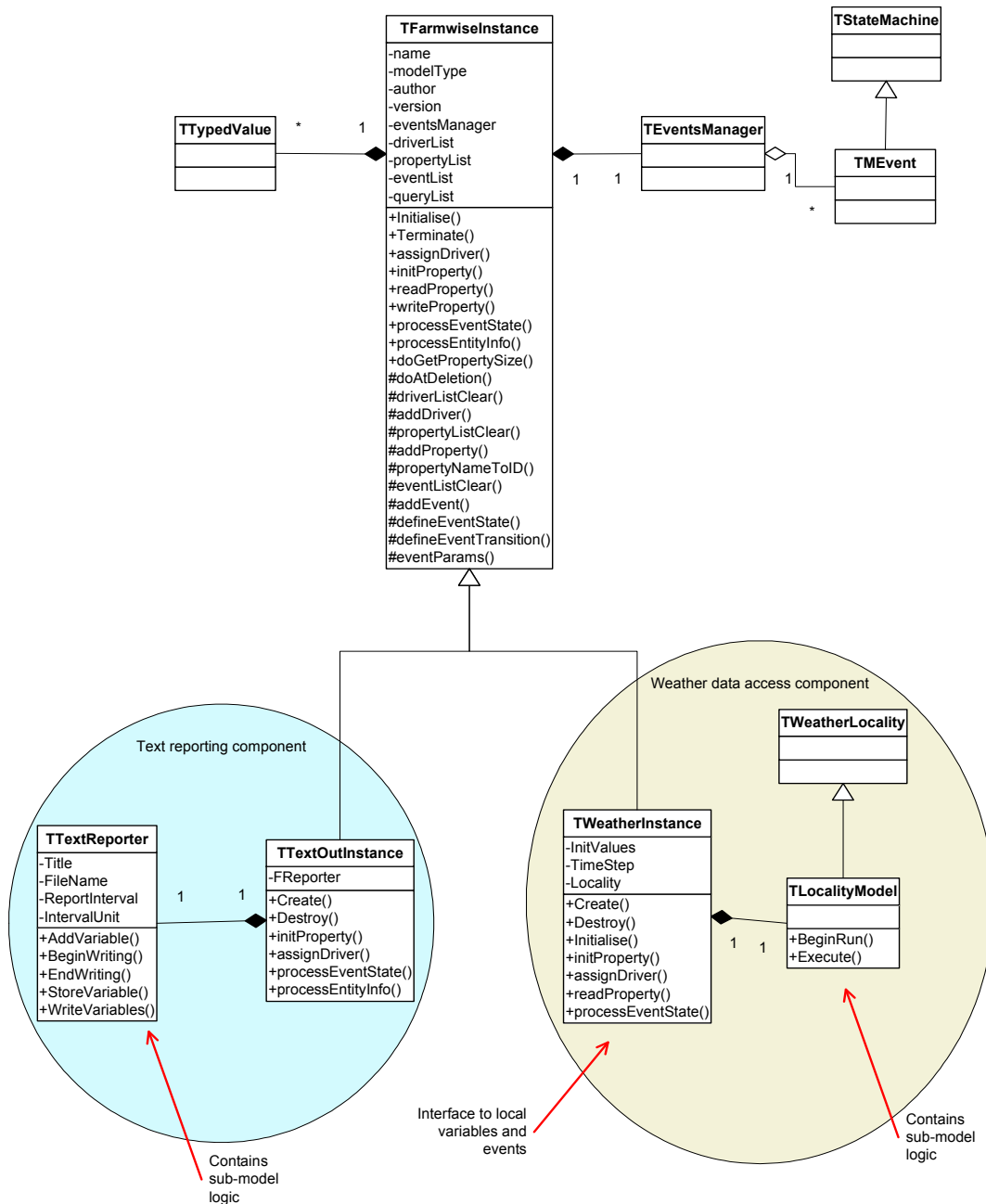


Figure 5.3. Classes involved in implementing two components by using *TFarmwiseInstance*.

Xerces <http://xml.apache.org/xerces-c/>
 XDOM <http://www.philo.de/xml/dom/>

5.3. Event processing using the state machine

The CMP implementation is designed to accommodate asynchronous processing. One consequence of this is that whenever an event handler generates a message that is sent to the rest of the simulation, the generating component must wait for the response to the message before executing the rest of the handler. In particular, this situation arises when the values of driving properties are requested: multiple properties can be requested at once, but the event handler must wait for all requests to be completed (and hence the driving values to be provided) before proceeding.

This requirement is met by implementing the code of each event handler within a *state machine*. The code of each event handler is divided into one or more *states*. The transitions from one state to another depend on the value of a *guard condition* that is set as each state is processed.

Figure 5.4 shows a common state machine for sequenced events, in which the values of driving properties are requested and then the event logic is computed. In this example, there is a single flow from state to state; in general different guard conditions can apply at the end of a state's processing, resulting in a state diagram with branching.

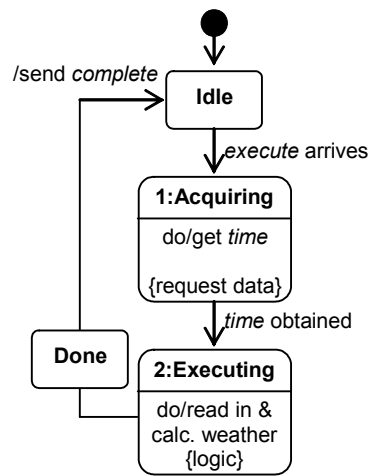


Figure 5.4. A simple state machine for an event handler.

6. Implementing a component: step by step

A component is implemented by creating a subclass from *TFarmwiseInstance* and overriding specific functions. When subclassing *TFarmwiseInstance*, the methods below are usually implemented:

Constructor	This is the place where component-specific properties and events are defined.
<i>initialise()</i>	Called twice during the creation of each module, once before and once after the module's initialisation information is passed to it. Pre- or post-processing of the initialisation information is carried out here.
<i>initProperty()</i>	Sets initial values for properties.
<i>readProperty()</i>	Used to provide the current values of owned properties to the rest of the simulation
<i>processEventState()</i>	Handles events that arrive from other parts of the simulation.

6.1. Defining the properties

Owned properties

In most cases, the properties and event handlers of a component are defined as part of the constructor for the *TFarmwiseInstance* descendant class. The *addProperty()* method is used to define each owned property's name, numeric ID, type and unit (where appropriate) and whether the property is readable, writable, and/or an initialisation property.

```
//addProperty(sName, ID, bRead, bWrite, bInit, sUnit, bIsArray, sType)
addProperty("start", prpSTART, false, false, true, "-", false, typeSTRING );
addProperty("finish", prpEND, false, false, true, "-", false, typeSTRING );
addProperty("day", prpDAY, true, false, false, "-", false, typeINT4 );
addProperty("month", prpMONTHNO, true, false, false, "-", false, typeINT4 );
addProperty("year", prpYEARNO, true, false, false, "-", false, typeINT4 );
```

- The *ID* parameter is an integer. It is good practice to use constant values for property and event IDs to ensure consistency between the methods that use them.
- The *sType* parameter is either a valid DDML type description, in which case the *sUnit* and *bArray* parameters are ignored, or it is one of a set of predefined constants that denote a scalar type, in which case the *sUnit* and *bArray* parameters are used. The most common scalar type parameters are `typeDOUBLE`, `typeINT4`, `typeSTRING` and `typeBOOLEAN`. A variety of other useful types are found in the *piTypes* code unit.
- Note that the first few property ID values are reserved for standard properties possessed by all components. Component-specific property ID values should be equal to or greater than the constant value `PROP_START_INDEX`.
- Initialisation code (see below) must be provided for all properties for which the *bInit* parameter is set to TRUE.

Driving properties

The *addDriver()* method is used to define each driving property's name, numeric ID and type, plus the valid range for the number of responses received to a request for the driver. The three type-related parameters are handled in the same way as for owned properties:

```
//addDriver( sName, ID, iMinConn, iMaxConn, sUnit, bIsArray, sType)
addDriver( 'shc_layer', drvSHC_L, 0, 1, '', FALSE, typeLAYERS_SHC );
addDriver( 'cover green', drvCOVER G, 0, NO_MAXCONN, '-', FALSE, typeDOUBLE );
addDriver( 'cover tot', drvCOVER T, 0, NO_MAXCONN, '-', FALSE, typeDOUBLE );
addDriver( 'sw_layer', drvSW_L, 1, 1, '', FALSE, typeLAYERS_SW );
```

6.2. Defining the event handlers

When defining each event handler, the parameters and the states and possible transitions in the state machine must be provided. The `addEvent()`, `defineEventState()` and `defineEventTransition()` methods are used for this purpose.

The following code defines the (sequenced) event handler shown in figure 5.4. A second, simpler event is also defined that does not require any driving property values; this state machine is typical for management events:

```
const
  evtEXEC    = 1;  evtSOMETHING = 2;
  smACQUIRE = 1;  smPROCESS    = 2;

  typePARAMS = '<type>'
              + '<field name="param1" kind="double" unit="kg/ha"/>'
              + '<field name="param2" kind="string"/>'
              + '</type>';
  :

// a simple sequenced event with driving property requests
addEvent( 'do process', evtEXEC, kindSUBSCRIBED, typeEMPTY ); // sName, ID, iKind, sType
defineEventState( evtEXEC, smACQUIRE, NONLOGIC ); // eventID, stateID, type
defineEventState( evtEXEC, smPROCESS, LOGIC );

defineEventTransition( evtEXEC, smACQUIRE, FAIL, DONE ); // eventID, stateID,
defineEventTransition( evtEXEC, smACQUIRE, 0, smPROCESS ); // condition, toState
defineEventTransition( evtEXEC, smPROCESS, FAIL, DONE );
defineEventTransition( evtEXEC, smPROCESS, 0, DONE );

// a simple management event
addEvent( 'do_something', evtSOMETHING, kindSUBSCRIBED, typePARAMS );
defineEventState( evtSOMETHING, smPROCESS, LOGIC );
defineEventTransition( evtEXEC, smPROCESS, 0, DONE );
```

- Note the use of `typeEMPTY` for `sType` in the first event; sequenced events may not have parameters as the sequencer component has no way of determining values for them.
- “Logic” states do not contain any communication with the rest of the simulation, and so processing can move to the next state immediately. “Non-logic” states must wait for information for the rest of the simulation before processing continues. A state should be defined as “non-logic” whenever it includes a property value request or a request for information about the structure of the simulation.

6.3. Initialising properties

When a simulation is executed, the initialisation process for each component results in three sets of method calls to the `TFarmwiseInstance` descendant:

1. The `initialise()` method is called with the `stage` parameter set to 1. Any component initialisation code that was not implemented in the constructor can be implemented here.
2. The `initProperty()` method is called for each initial value provided in the simulation’s SDML document. The implementation of `initProperty()` must store these value data, often by passing them to the model logic object.
3. Once all the initial value data have been passed, the `initialise()` method is called a second time with `stage` parameter set to 2. Any component initialisation code that depends on combining initialisation values can be implemented here.

As a result, the *initialise()* and *initProperty()* implementations tend to look like this:

```
procedure TMyInstance.Initialise( iStage : Integer );
begin
  if (iStage = 1) then
    begin
      /** pre-processing code here **//
    end
  else if (iStage = 2) then
    begin
      /** combine any initialisation values here **//
    end
  end;

procedure TMyInstance.initProperty( iPropertyID : Integer; aValue : TTypedValue );
begin
  case iPropertyID of
    prpSTART : FModel.setStart( aValue.asString );
    prpEND   : FModel.setEnd(   aValue.AsString );
    // etc

    else raise Exception.Create( 'Invalid ID code in initProperty' );
  end;
end;
```

6.4. Acquiring values for driving properties

Driving properties must be defined before they can be requested (see section 6.1). Values of driving properties are requested within event handlers by calling the *sendDriverRequest()* method. Zero or more responses to the request may be received. A call to the *assignDriver()* method will be made for each valid response; this method must therefore implement code to store or otherwise handle the driving variable's value.

A typical implementation of *assignDriver()* looks like this:

```
procedure TMyInstance.assignDriver( iDriverID : longword;
                                   iProvider : Integer;
                                   aValue     : TTypedValue );
begin
  case iDriverID of
    drvCOVER_G : FCoverGreenAll := 1.0 - (1.0-FCoverGreenAll) * (1.0-aValue.asDouble);
    drvCOVER_T : FCoverTotalAll := 1.0 - (1.0-FCoverTotalAll) + (1.0-aValue.asDouble);
    drvSW L    : Layers2LayerArray( aValue, FTheta );
    // etc

    else
      inherited assignDriver(iDriverID, iProvider, aValue);
  end;
end;
```

6.5. Implementing model logic – event handlers

All event-handling code must be implemented within the *processEventState()* function of the *TFarmwiseInstance* descendant class. Code must be written for each state of each event. The function result must give a valid guard condition value for each state of each event (i.e. one that has been given in a call to *defineEventTransition()*) as it will determine the next transition of the state machine.

The following function might implement the two events that were defined in the example code in section 6.2:

```

{=====}
{ processEventState
{ Selects the event and its state to be processed.
{ Returns: The guard condition/result of the state's processing
{=====}

function TMyInstance.processEventState( iEventID, iState : Integer;
                                       iSender      : LongWord;
                                       Params       : TTypedValue): integer;

var
  iCondition: integer; //guard condition to return
begin
  // do process event

  if (iEventID = evtEXEC) then
  begin
    case iState of //choose the state to execute
      smACQUIRE : begin
        sendDriverRequest( drvTIME, iEventID ); // causes FTimeStep to be
          iCondition := 0; // populated
        end;
      smPROCESS : begin
        FModel.Execute( FTimeStep.Start.iDay );
          iCondition := 0;
        end;

      else iCondition := inherited processEventState(iEventID,iState,iSender,Params);
    end;
  end

  // do something event

  else if (iEventID = evtSOMETHING) and (iState = smPROCESS) then
  begin
    FModel.doSomething( Params.Member['param1'].asDouble,
                      Params.Member['param2'].asString );

    iCondition := 0
  end

  else
    iCondition := inherited processEventState( iEventID, iState, iSender, Params );

  Result := iCondition;
end;

```

Note the way in which:

- each state of each event has a block of code that sets the guard condition;
- information from the rest of the simulation (driving property values or event parameters) is passed to methods of the model logic class for processing;
- all event-state combinations that were not defined as part of the event's state machine are passed to the inherited version of *processEventState()*

6.6. Providing values for owned properties

Requests from the rest of the simulation for the value of an owned property will result in a call to the *readProperty()* method of the *TFarmwiseInstance* descendant class. Code must be written to populate the *TTypedValue* parameter of this method with the current value of the nominated property. The value parameter is guaranteed to be of the correct type, but the lengths of any arrays in the value must be set to the correct length as part of populating the value.

The `readProperty()` method will typically look something like this:

```
procedure TMyInstance.readProperty(  
    iPropertyID : Integer          // local ID of owned property  
    iRequestor  : Integer;        // ID of requesting component  
    aValue      : TPropertyInfo ); // TTypedValue descendant  
begin  
    case iPropertyID of  
        prpELEV  : aValue.setValue( FModel.fElevationM );           // populate scalar values  
        prpLATT  : aValue.setValue( FModel.fLatDegrees );  
        prpLONG  : aValue.setValue( FModel.fLongDegrees );  
        prpMAXT  : aValue.setValue( FModel.getData( wdtMaxT ) );  
        prpMINT  : aValue.setValue( FModel.getData( wdtMinT ) );  
        prpWEATHER : begin // populate a record  
            aValue.Item[1].setValue( FModel.getData( wdtMaxT ) );  
            aValue.Item[2].setValue( FModel.getData( wdtMinT ) );  
            aValue.Item[3].setValue( FModel.getData( wdtRain ) );  
            aValue.Item[4].setValue( FModel.getData( wdtSnow ) );  
            aValue.Item[5].setValue( FModel.getData( wdtRadn ) );  
            aValue.Item[6].setValue( FModel.fVP_Deficit );  
            aValue.Item[7].setValue( FModel.getData( wdtWind ) );  
        end;  
        // etc  
    else raise Exception.Create( 'Invalid ID code in readProperty' );  
    end;  
end;
```

6.7. Publishing events

The process for sending an event from a component to the rest of the simulation (thereby triggering a computation in another component or components) is:

1. Define the name and type of the event to be published via a call to `defineEvent()` with the *kind* parameter set to `kindPUBLISHED`. This is usually done in the subclass constructor.
2. Call the `sendPublishEvent()` when the event is to be triggered. This function takes the numeric ID of the event as an argument. If the event has parameters, they reside in a *TTypedValue* that can be obtained via a call to the `eventParams()` method; they must be populated before the call to `sendPublishEvent()`.

```
publParams := eventParams( 33 ); // eventID  
  
publParams.Member['param1'].asString := 'something';  
publParams.Member['param1'].asDouble := 0.0;  
  
sendPublishEvent( 33, TRUE ); // eventID, bAcknowledge
```

For further discussion of published events, see section 2.3 of the protocol specification document.

6.8. Handling errors

Components may report errors at any time during the simulation, or while the simulation is starting or terminating. Errors fall into two categories; fatal and non-fatal. Fatal errors will cause the simulation to terminate. Within a *TFarmwiseInstance* subclass, the `sendError()` method is used to report an error and, if necessary, to terminate the simulation. As a rule, therefore, exception-handling code should include a call to `sendError()`.

```
try  
    // processing code here  
except  
    on E : Exception do  
        sendError( 'Error in my_comp: '+E.Message, TRUE );  
end;
```

6.9. Discovering the current time step

To discover the start and end times of the current time step, a component must define the `time` driving property. This property must have the `typeTIMESTEP` type, which is a record with the following fields:

```
<type>
  <field name="startDay"      kind="integer4" unit="d" />
  <field name="startSec"     kind="integer4" unit="s" />
  <field name="startSecPart" kind="double"   unit="s" />
  <field name="endDay"       kind="integer4" unit="d" />
  <field name="endSec"       kind="integer4" unit="s" />
  <field name="endSecPart"   kind="double"   unit="s" />
</type>
```

The two day fields in this type are stored as Julian Day Numbers. *TTypedValue* data for this property can be decoded and manipulated using the *TTimeValue* class that is provided as part of the component API.

7. Advanced topics

7.1. The connection scheme

When a simulation contains multiple modules drawn from the same component, the question arises: which of many possible sources for a driving property should be used? We refer to the rules for determining this as a *connection scheme*.

In the CPI protocol implementation, connections for a driving property are made to those source components that are equally-nearest to the requesting component. “Distance” is calculated using the number of steps up and down the hierarchical tree of components. Components cannot connect to themselves.

Whenever the simulation is restructured, the system components update the set of connections for each driving property belonging to components within their system.

7.2. Querying the simulation for its structure

The *sendQueryInfo()* method requests information about a component, property or event elsewhere in the simulation. Responses will result in a call to the *processEntityInfo()* method.

7.3. Other TFarmwiseInstance methods that generate messages

<i>sendPauseSim()</i>	Pauses or resumes a simulation.
<i>sendActivateComp()</i>	Activates another component.
<i>sendEndSimulation()</i>	Terminates the simulation.
<i>sendComplete()</i>	Sends an acknowledgement to another component. This method should only be called in a context mandated by the protocol specification.
<i>addSetterProperty()</i>	Adds a property reset request to the component and registers it.
<i>sendReqSetValue()</i>	Requests that the value of another component’s property be reset.
<i>deleteProperty()</i>	Removes an owned property from the component. It is then deregistered.
<i>deleteDriver()</i>	Removes a driving property from the component. It is then deregistered.
<i>deleteSetterProperty()</i>	Removes a property reset request from the component. It is then deregistered.
<i>deleteEvent()</i>	Removes an event from the component and deregisters it.

8. Further documentation

8.1. Common Modelling Protocol specification

The Common Modelling Protocol has been defined in detail in a public document entitled: ‘Specification of the CSIRO Common Modelling Protocol’. This document covers in detail; definitions, system tasks, messaging, properties and events, data types, SDML, and some implementation techniques.

8.2. Typed Value specification

Throughout the PI implementation of the CMP, *TTypedValue* and its descendant classes are used to contain structured types with their definitions and values. This technique allows for types of any complexity to be specified and manipulated.

The document “Typed Value objects in the CPI implementation of the Common Modelling Protocol.” contains details about the *TTypedValue* class.

8.3. Help file

A Windows Help file named CMP_HELP.CHM accompanies this document. This file contains more detailed documentation of the classes used in the component-building API.