A Formal Approach for Designing
CORBA based Applications§ ¥

Alberto Coen-Porisini♣
Matteo Pradella♣
Matteo Rossi♣
Dino Mandrioli♣

(♣) Dipartimento di Elettronica e Informazione - Politecnico di Milano
Piazza L. da Vinci, 32 – I-20133 Milano, Italy
ph: +39-02-2399-{3561, 3522}, email: {rossi, mandrioli}@elet.polimi.it

(♣) Dipartimento di Informatica e Comunicazione - Università dell’Insubria
Via Ravasi, 2 – I-21100 Varese, Italy
ph: +39-0332-218943, email: alberto.coenporisini@uninsubria.it

(♣) CNR Istituto di Elettronica e di Ingegneria dell’Informazione e delle Telecomunicazioni -
Sez. Milano
Via Ponzio 34/5 – I-20133 Milano, Italy
ph: +39-02-2399-3495, email: pradella@elet.polimi.it

ABSTRACT
The design of distributed applications in a CORBA based environment can be carried out by
means of an incremental approach, which starts from the specification and leads to the high
level architectural design. This paper discusses a methodology to transform a formal
specification written in TRIO into a high level design document written in an extension of
TRIO, named TRIO/CORBA (TC). The TC language is suited to formally describe the high
level architecture of a CORBA based application. As a result, designers are offered with high-
level concepts that precisely define the architectural elements of an application. Furthermore,
TC offers mechanisms to extend its base semantics, and can be adapted to future
developments and enhancements in the CORBA standard. The methodology and the
associated language are presented by means of a case study derived from a real Supervision
and Control System.

Keywords
CORBA, Architectural Design, Object Orientation, Frameworks, Formal Methods, Temporal
Logic, Supervision and Control Systems

1 INTRODUCTION
Application development is composed of three major phases: requirement analysis and
specification, architectural design, implementation, whose correctness must be rigorously
verified. Great benefits in terms of user requirements validation and verification of the
implemented system can be obtained if the specification is expressed in a precise (possibly
formal) way, and if the designer is supported by a methodology (and related tools) for deriving
the architecture of the application from the specification.

§ This work has been partially supported by the Commission of the European Union – ESPRIT Project
OpenDREAMS-II and by the Italian Ministero dell’Università e della Ricerca Scientifica – Progetto Mosaico
¥ This paper is an expanded version of the paper by [Coen-Porisini et al. 2000].
Formal methods, although their wide adoption is rather controversial in industrial environments, are often strongly recommended to enhance reliability of critical systems such as Supervision and Control Systems (SCSs). Despite the still ongoing debate, a fairly large amount of experience has demonstrated the effectiveness of formal methods at least in the realm of safety critical, possibly real-time, applications [Hinchey and Bowen 1995, Saiedain et al. 1996, Ciapessoni et al. 1999].

SCSs are often physically distributed over local or wide area networks and are usually implemented as closed systems based on proprietary hardware and software; thus, they are usually not portable and cannot be extended or integrated into more complex systems. Therefore, adding new functionalities to existing SCSs often leads to building new independent systems. For instance, an Energy Management System is typically composed of several independent applications, each of which has its own sensors, hardware processors, databases and specialized software, although conceptually they share the same information. Since the functional architecture of all these applications is very similar, several components are duplicated (e.g., there is a data acquisition component for each application).

On the other hand, open environments, where different applications can coexist and share information, are gaining more and more acceptance to overcome the typical difficulties of distributed, heterogeneous, often legacy, systems. One promising possibility is to use the Common Object Request Broker Architecture (CORBA) technology\(^1\) [OMG 2002a]. In fact, the Object Management Group (OMG) has defined the Object Management Architecture [Soley and Stone 1995], which addresses both general issues and particular needs of specific application domains (e.g., Banking, Telecom, Supervision and Control Systems) by defining high level libraries or \textit{frameworks} [Fayad et al. 1999].

CORBA supports the extension of an SCS by adding new components whenever they are developed, thus reducing development time and cost. For instance, alarms could be recorded by the alarm managing subsystems and accessed through a global database by the diagnostic subsystem. To fully achieve such a goal, however, two crucial issues must be addressed:

- While some issues that are critical for SCSs (such as reliability and real-time) have been taken into account by the OMG since 2000 [OMG 2000, OMG 2002a, OMG 2002b], still only a handful of CORBA implementations exist (e.g., TAO [Schmidt], VisiBroker-RT [Borland]) that are (sometimes not fully) compliant to the new standard.
- A big gap must be filled up by design to move from application requirements to a complete CORBA-based implementation.

This paper addresses the latter issue by presenting an approach for designing distributed systems in a CORBA environment, based on an initial formalization of the requirements given in terms of TRIO [Ghezzi et al. 1990, Morzenti and San Pietro 1994]. TRIO is a first order temporal logic which has shown to be effective for analyzing and specifying critical systems, such as SCSs [Ciapessoni et al. 1999].

In general, TRIO-based methods, including the one presented in this paper, can be classified as “lightweight formal methods” [Saiedain et al. 1996, Easterbrook et al. 1998] as they allow \textit{but not enforce} the use of formalisms, whenever designers decide that their benefits are worth the related cost. In other words, they can be effective even if they are employed (possibly in conjunction with other, informal notations) to describe/validate/verify only those parts of the system where precision and formality are crucial, while the rest of the application can be informally defined. For instance, only critical requirements could be formalized and the

---

\(^1\) In what follows we assume the reader has some knowledge of CORBA concepts and terms [Siegel 2000]
implementation could possibly be proved mathematically correct only against (some of) them, leaving less critical aspects for informal and less expensive treatment.

This paper presents a methodology that consists of moving from the TRIO specification of application requirements to a new formalization representing the high level architectural design in which the technological target (i.e., CORBA) is taken into account. This transformation is supported by a language, whose name is TC (TRIO/CORBA), obtained by introducing and formalizing in TRIO the basic concepts characterizing CORBA. The integration of a formal approach during the specification phase with CORBA concepts, at the design level, is expected to enhance the development process.

TC combines a flexible formal notation that explicitly deals with time with a methodology to design the high level architecture of CORBA-based systems. As a result, TC enjoys the benefits of formality (precision, verifiability, etc.), while designers are offered with high-level concepts that precisely define the architectural elements of an application. Furthermore, TC offers mechanisms to extend its base semantics, and can be adapted to future developments and enhancements in the CORBA standard, and possibly applied to other middleware technologies.

The example presented in this paper refers to a SCS, namely an Energy Management System, that has been specified and then designed to run in a CORBA-based environment. Although the chosen example presents several critical (timing) requirements, the paper focuses on the architectural language (TC) and the associated methodology used to design such systems, rather than on the way in which such critical requirements are specified and verified\(^2\). We believe that the results are general enough to be applied to critical applications in almost any domain. Moreover, both the language and the methodology are used to take into account also non-critical requirements, even though whether this is cost-effective is still under debate.

The results reported in this paper are part of a long-term research effort carried on by our group at Politecnico di Milano in cooperation with several academic and industrial partners. A major part of such a research was accomplished within the ESPRIT projects OpenDREAMS I and II\(^3\). Such projects aimed at building a complete CORBA-based platform suitable for the development of SCSs; the platform includes a whole suite of CORBA services and frameworks and of development-supporting tools. The technical documentation of the projects [OpenDREAMS, OpenDREAMS-II 2000] provides a complete description of the achievements.

The paper is organized as follows: Section 2 provides a short summary of TRIO; Section 3 discusses the main features of TC; Section 4 presents the methodology by means of a case study in which TC is used to design a Supervision and Control System; Section 5 discusses some of the benefits of our formal approach, with an emphasis on the lessons learned from applying the TC methodology to the case study outlined in Section 4; Section 6 compares the research presented in this paper with some related works; finally Section 7 draws some conclusions and outlines future works.

### 2 A SUMMARY OF THE TRIO SPECIFICATION LANGUAGE

TRIO [Ghezzi et al. 1990, Morzenti and San Pietro 1994] is a first order temporal logic language that supports a linear notion of time. Besides the usual propositional operators and quantifiers, one may compose formulas by using a single basic modal operator, called Dist, that relates the current time, which is left implicit in the formula, to another time instant: the

---

2 These issues are the objects of companion reports [Marotta 2001, Rossi 2002].

3 OpenDREAMS-I started in 1995 and lasted one year, while OpenDREAMS-II started in 1997 and ended in May 2000
formula $\text{Dist}(F, t)$, where $F$ is a formula and $t$ a term indicating a time distance, specifies that $F$ holds at a time instant at $t$ time units from the current instant.

A number of derived temporal operators can be defined from the basic $\text{Dist}$ operator through propositional composition and first order quantification on variables representing a time distance. Table 1 reports the formal definition of some TRIO derived operators along with a short informal description of their semantics. Most of such operators are symmetrically defined with reference to the past and the future of the current instant.

<table>
<thead>
<tr>
<th>Operator</th>
<th>Definition</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Past(A, d)</td>
<td>$d &gt; 0 \land \text{Dist}(A, -d)$</td>
<td>A held (holds) $d$ time units in the past (future)</td>
</tr>
<tr>
<td>Futr(A, d)</td>
<td>$d &gt; 0 \land \text{Dist}(A, d)$</td>
<td>A held (holds) in the future over a period of $d$ time units</td>
</tr>
<tr>
<td>SomP(A)</td>
<td>$\exists t (t &gt; 0 \land \text{Dist}(A, -t))$</td>
<td>A held (holds) sometimes in the past (future), i.e., there is a past (future) instant in which $A$ held (holds)</td>
</tr>
<tr>
<td>SomF(A)</td>
<td>$\exists t (t &gt; 0 \land \text{Dist}(A, t))$</td>
<td>A held (holds) in the past (future) over a period of $d$ time units</td>
</tr>
<tr>
<td>WithinP(A, d)</td>
<td>$\exists t (0 &lt; t &lt; d \land \text{Dist}(A, -t))$</td>
<td>A occurred (occurs) in an instant at most $d$ time units in the past (future)</td>
</tr>
<tr>
<td>WithinF(A, d)</td>
<td>$\exists t (0 &lt; t &lt; d \land \text{Dist}(A, t))$</td>
<td>A occurred (occurs) in an instant at most $d$ time units in the future (past)</td>
</tr>
<tr>
<td>Lasted(A, d)</td>
<td>$\forall t (0 &lt; t &lt; d \rightarrow \text{Dist}(A, -t))$</td>
<td>A held (holds) in the past (future) over a period of $d$ time units</td>
</tr>
<tr>
<td>Lasts(A, d)</td>
<td>$\forall t (0 &lt; t &lt; d \rightarrow \text{Dist}(A, t))$</td>
<td>A held (holds) in the past (future) over a period of $d$ time units</td>
</tr>
<tr>
<td>Since(A, B)</td>
<td>$\exists t (t &gt; 0 \land \text{Dist}(B, -t) \land \text{Lasted}(A, t))$</td>
<td>There is a past (future) instant in which $B$ held (holds), and $A$ was (is) true since (until) that moment</td>
</tr>
<tr>
<td>Until(A, B)</td>
<td>$\exists t (t &gt; 0 \land \text{Dist}(B, t) \land \text{Lasts}(A, t))$</td>
<td>There is a past (future) interval of non null length (starting from the current instant) in which $A$ held (holds)</td>
</tr>
<tr>
<td>UpToNow(A)</td>
<td>$\exists t (t &gt; 0 \land \text{Lasted}(A, t))$</td>
<td>There is a past (future) interval of non null length (starting from the current instant) in which $A$ held (holds)</td>
</tr>
<tr>
<td>NowOn(A)</td>
<td>$\exists t (t &gt; 0 \land \text{Lasts}(A, t))$</td>
<td>There is a past (future) interval of non null length (starting from the current instant) in which $A$ held (holds)</td>
</tr>
<tr>
<td>LastTime(A, d)</td>
<td>$d \geq 0 \land \text{Dist}(A, -d) \land \text{Lasted}(\neg A, d)$</td>
<td>The last (next) time that $A$ held (holds) was (will be) $d$ time units ago (in the future)</td>
</tr>
<tr>
<td>NextTime(A, d)</td>
<td>$d \geq 0 \land \text{Dist}(A, d) \land \text{Lasts}(\neg A, d)$</td>
<td>The last (next) time that $A$ held (holds) was (will be) $d$ time units ago (in the future)</td>
</tr>
</tbody>
</table>

Table 1: Some TRIO derived operators

TRIO is well suited to deal with both continuous and discrete time. What changes in the two cases is the domain over which time variables range (real or integer numbers). In this paper the time domain is assumed to be continuous.

The natural tendency to describe systems in an operational way is supported by TRIO through the so-called ontological constructs, such as events and states. An event is a particular predicate that is supposed to model instantaneous conditions such as a change of state or the occurrence of an external stimulus. Events can be associated with conditions that are related causally or temporally with them. A state is a predicate representing a property of a system. A state may have duration over a time interval; state changes may be associated with suitable pre-defined events and conditions.

The semantics of such constructs is defined by means of (pre-defined) TRIO axioms. In fact a distinguishing feature of TRIO is that every high level concept is defined in terms of lower level ones down to the operator $\text{Dist}$. For example, event $E$ must satisfy the following (non-Zeno\(^4\)) behavior (as defined in [Gargantini and Morzenti 2001]):

$$\text{UpToNow}(\neg E) \land \text{NowOn}(\neg E)$$

\(^4\) An event has non-Zeno behavior if it cannot occur infinitely many times in a finite interval (i.e., it does not have accumulation points).
TRIO items (values, predicates, functions, events, states, etc.) are distinguished into time-independent (TI), whose value does not change during system evolution, and time-dependent (TD), whose value may change during system evolution.

For specifying large and complex systems, TRIO has the usual object-oriented concepts and constructs such as classes, inheritance and genericity. Classes denote collections of objects (class instances) that satisfy a set of axioms. Notice that TRIO, being a logic language, does not support object creation/ destruction. Therefore, if one wants to model an entity having a limited lifetime he/she must simulate creation/destruction using other TRIO mechanisms, such as a time-dependent predicate that is true when the object exists, and false otherwise. In addition, TRIO objects do not have a priori a unique identifier to distinguish one object from the others. However, object identity can be modeled by introducing an item that represents the identity of the object and some axioms assuring that different objects have different identities.

Classes can be either simple or structured –the latter term denoting classes obtained by composing simpler ones. A simple class is defined through a set of axioms premised by a declaration of all items that are referred therein. Some of such items are visible, that is, they may be referenced from outside, in the context of a complex class whose instances include an instance of that class.

For example, let us consider a system composed of a user-operated console that acquires measurements made by two identical sensors and computes the average value. When the user pushes a “start/stop” button, the console starts/stops sending the computed value to some output device. Furthermore, the console can be turned on/off by pressing another button.

The formalization of the console is given by class \texttt{Console\_Class} that includes two events (\texttt{Toggle\_start/stop} and \texttt{Switch\_on/off}), one state (\texttt{Console\_On}) and three values (\texttt{Sensor1\_measurement}, \texttt{Sensor2\_measurement} and \texttt{Computed\_measurement}). The semantics is given by axioms describing the behavior of consoles. For instance, a first axiom states that \texttt{Console\_On} is true in the current time instant if and only if there exists a time instant in the past when the console was on, and it has not been turned off since then, by pressing the on/off button; in other words, \texttt{Console\_On} is switched on unless the button has been pressed:

\[ \texttt{Console\_On} \leftrightarrow \text{Since}(\neg \texttt{Switch\_on/off}, \texttt{Console\_On}) \]

A similar axiom relates the \texttt{Toggle\_start/stop} button with the sending of measurements from the sensors; and a third axiom states that \texttt{Computed\_measurement} is the average of the values \texttt{Sensor1\_measurement} and \texttt{Sensor2\_measurement}. Therefore, the definition of class \texttt{Console\_Class} is as follows:

\begin{verbatim}
  Class Console\_Class
    visible  Sensor1\_measurement, Sensor2\_measurement, Computed\_measurement,
             Toggle\_start/stop, Switch\_on/off
    temporal domain real

  TD Items
    value Sensor1\_measurement : real;
    value Sensor2\_measurement : real;
    value Computed\_measurement : real;
    state Console\_On;
    event Toggle\_start/stop;
    event Switch\_on/off;

  axioms
    [ ... axioms that define the behavior of the class ... ]

end Console\_Class
\end{verbatim}
TRIO is also endowed with a graphic representation in terms of boxes, lines, and connections to depict class instances and their components, information exchange, and logical equivalence among (parts of) objects. For example, Figure 1 shows the graphic description of Console_Class. Visible items cross the box, that is, they may be referenced from outside, in the context of a complex class whose instances include an instance of Console_Class.

![Figure 1: A TRIO simple class](image)

The other components of the system can be specified in a similar way, leading to the definition of classes Output_device_Class, Console_User_Class and Sensor_Class to represent the output device, the user operating the console and the sensors from which data are acquired, respectively.

The whole system is formalized by means of structured class Overall_System, whose instances contain one instance of Console_Class, Output_device_Class and Console_User_Class and two instances of Sensor_Class. In a structured class, the classes composing it are also called modules (see declaration of Overall_System below); an instance of a structured class contains an instance for each of its modules. Figure 2 depicts class Overall_System.

![Figure 2: A TRIO structured class](image)

A line (called TRIO connection) between two items means that they are logically equivalent and therefore it does not imply any direction in the flow of information. If connected items share the same name, this is written outside the classes that contain it, and the corresponding
The textual definition of class Overall_System is the following:

Class Overall_System
temporal domain real

modules
  Console : Console_Class;
  Output_device : Output_device_Class;
  Sensor1 : Sensor_Class;
  Sensor2 : Sensor_Class;
  User : Console_User_Class

connections
  (connect Output_device.Switch_on/off, User.Switch_Dev_on/off)
  (connect Console.Toggle_start/stop, User.Toggle_start/stop)
  (connect Console.Computed_measurement, Output_device.Input_value)
  (connect Console.Sensor1_measurement, Sensor1.measurement)
  (connect Console.Sensor2_measurement, Sensor2.measurement)
[ ... rest of connections not shown ... ]

axioms
[ ... axioms that define the behavior of the class ... ]

end Overall_System

Notice that the specification of any non-trivial system is usually made up of one structured class, which models the overall system along with its environment, whose instances include instances of other classes representing the different components of the system. The global semantics of a structured class is defined by the logical conjunction of all axioms of the class and of its modules.

3 THE TC LANGUAGE

The TRIO/CORBA (TC) language enriches TRIO with the typical elements of CORBA, allowing designers to rigorously describe the architecture of a CORBA application. TC has the formal rigor of TRIO and is suitable for describing the high level design of an application by allowing one to formally define the behavior of the components of an architecture and the way in which they interact.

TC introduces all basic CORBA concepts such as operations, attributes, exceptions, interfaces, and so on. Then, complex concepts (e.g., services, frameworks) are built from such basic elements. Following the typical TRIO approach, every basic concept has a semantics, which is formalized by means of (pre-defined) axioms. The collection of these axioms describes in an abstract way the behavior of any CORBA-based system. In this way designers can focus on higher-level user-defined requirements, while low-level aspects can remain hidden. Notice that TC has been extended to include also the low level aspects of any CORBA implementation (e.g., the Portable Object Adapter), thus leading to a two-layered description of CORBA\(^5\). This paper focuses on the main aspects of the first (most abstract) layer, while the interested reader can refer to [Rossi 2002], where both layers are presented in a more comprehensive way.

---

\(^5\) The need for the second layer usually arises when moving from design to verification.
In order to formalize the basic CORBA concepts TC introduces four “meta-classes”: Interface, CORBA Entity, NonCORBA Entity and Environment\(^6\). A TC meta-class can be viewed as a template used for generating TRIO classes that share a set of common properties described by means of (pre-defined) axioms. Thus, an instance of a TC meta-class is a TRIO class, while an instance of a TRIO class is a TRIO object.

Interface and CORBA Entity meta-classes describe the properties of CORBA IDL interfaces and of objects interacting with an Object Request Broker (ORB), respectively. NonCORBA Entity meta-class models non-CORBA-related entities, while the Environment meta-class is used to structure the description of an architecture in terms of the above mentioned meta-classes.

In the rest of the paper the following convention is adopted: CORBA Entity denotes the name of a TC meta-class while CORBA Entity Class C denotes a class named C, instance of the meta-class CORBA Entity. For the sake of readability whenever no ambiguity can arise, we refer to a CORBA Entity Class C as CORBA Entity C. The same convention is also used for the instances of the other meta-classes.

In what follows, the main features of the different TC meta-classes are discussed with reference to CORBA 2.5 specification [OMG 2001a].

### 3.1 The Interface meta-class

According to OMG’s definition, a CORBA IDL interface defines the signature of a set of operations/attributes that must be provided by any CORBA object satisfying such an interface. Therefore, an Interface class introduces only the signature of its operations/attributes without providing their semantics\(^7\). Thus, no axiom is defined in an Interface class.

For example, Interface TerminalInterface introduces the user-defined type IntReturnedType, the exception negativeOrZeroPar and operations getString and getInt. Operation getString takes one input parameter (maxLength of type short integer), returns a value of type string, and can raise an exception (negativeOrZeroPar), while operation getInt does not take any input parameter and does not raise any exception, and returns a value of type IntReturnedType.

<table>
<thead>
<tr>
<th>Interface Class</th>
<th>TerminalInterface</th>
</tr>
</thead>
<tbody>
<tr>
<td>type</td>
<td>IntReturnedType = long;</td>
</tr>
<tr>
<td>exceptions</td>
<td>negativeOrZeroPar;</td>
</tr>
<tr>
<td>operations</td>
<td>getString</td>
</tr>
<tr>
<td></td>
<td>parameters</td>
</tr>
<tr>
<td></td>
<td>in maxLength: short;</td>
</tr>
<tr>
<td></td>
<td>returns string;</td>
</tr>
<tr>
<td></td>
<td>raises negativeOrZeroPar;</td>
</tr>
<tr>
<td></td>
<td>getInt</td>
</tr>
<tr>
<td></td>
<td>returns IntReturnedType;</td>
</tr>
<tr>
<td>end TerminalInterface</td>
<td></td>
</tr>
</tbody>
</table>

TC declaration of Interface class TerminalInterface

```plaintext
interface TerminalInterface {
    typedef long IntReturnedType;

    exception negativeOrZeroPar();

    string getString (in short maxLength)
        raises (negativeOrZeroPar);

    IntReturnedType getInt();
}
```

Corresponding IDL interface declaration

---

\(^6\) Meta-classes are denoted in courier font.

\(^7\) For the sake of precision, an IDL interface provides some information about the semantics of the method calls (oneway or synchronous). However, in TC such aspect is modeled by means of stereotypes, which are introduced in Section 3.6, and thus a TC Interface has no semantics.
3.2 The CORBA Entity meta-class

The CORBA Entity meta-class formalizes the features of any component interacting with an ORB. According to OMG's terminology, entities providing operations that can be invoked through the ORB are referred to as CORBA objects or servers, while entities invoking operations of a CORBA object through an ORB are referred to as clients. Notice that a server can invoke operations provided by another server and thus in this case it can also be viewed as a client. An entity invoking operations provided by servers and not providing any operation is referred to as a pure client. The OMG requires that every CORBA object satisfies at least one IDL interface, while a pure client need not satisfy any interface.

In the OMG/CORBA specification, the definition of the satisfies relationship between servers and IDL interfaces depends on the implementation language. For example, in C++ the class implementing the server inherits from the virtual class representing the IDL interface. In Java the class implementing the server implements the Java interface representing the IDL interface. Finally, in C the file containing the implementation of the server includes the file containing the prototypes representing the IDL interface.

TC models the "satisfies" relationship between an IDL interface and a server by means of inheritance. In fact, a CORBA Entity class, whose instances represent a server, inherits from the Interface class modeling the IDL interface. Instead, a CORBA Entity class modeling a pure client does not inherit from any Interface class.

Since an Interface class cannot contain any axiom, different CORBA Entity classes can be designed to provide different semantics to the same Interface class according to the definition of IDL interface.

In what follows the term object refers to TRIO objects, that is, instances of TRIO classes, while the term server refers to CORBA objects. Thus, an instance of a CORBA Entity class is a TRIO object representing either a server (CORBA object) or a client. For the sake of readability, we refer to a CORBA Entity class whose instances are servers (clients) as a server class (client class). Notice that a class C can be both a server and a client class. Moreover, given a server class S satisfying an IDL interface I containing operation Op, S is said to export Op. Conversely, given a client class C invoking the operation Op, C is said to use I and to import Op.

All CORBA Entity classes share a set of properties, expressed by means of axioms, common to all entities that can interact by means of an ORB. For instance, they have an item _id that is used to uniquely identify every instance of a CORBA Entity class:

\[ _id : OID \]

OID is a TC basic type representing the set of all possible identifiers that can be assigned to an instance of a CORBA Entity class.

Notice that _id is an abstraction that can be used to model the object identity of any CORBA object (server) as defined by the IdentifiableObject interface of the CORBA Relationship service. Let us consider a CORBA object O whose item _id evaluates to val_id, then val_id represents the identity of O. Moreover, The attribute _id can also be viewed as the oid attribute that can be established through the CORBA Portable Object Adapter (POA). POAs, are not introduced in the paper since their existence can be hidden to designers who do not need to focus on the way in which CORBA is actually implemented.

Given a server class exporting operation returnType Op(a1, ..., an), the following TRIO events are introduced:
- \textit{Op(i).inv\_received}, is true when the server\(^8\) exporting \textit{Op} receives invocation \textit{i}\(^9\) of \textit{Op};
- \textit{Op(i).start} is true when invocation \textit{i} starts to be processed by the server;
- \textit{Op(i).end\_success} is true when invocation \textit{i} ends without an exception;
- \textit{Op(i).Exc.raise} is true when, on the server, invocation \textit{i} raises exception \textit{Exc} (standard or user-defined).

Moreover, given a client class importing operation \textit{Op}, the following events are defined:
- \textit{Op(i).invoke} is true when the client issues invocation \textit{i} of operation \textit{Op};
- \textit{Op(i).reply} is true when the client receives the reply from invocation \textit{i} of \textit{Op} and no exception is raised (either by the server or the ORB);
- \textit{Op(i).Exc.received} is true when invocation \textit{i} of operation \textit{Op} terminates on the client side with exception \textit{Exc} (standard or user-defined); a client receives an exception when either the server, or the ORB raises it.

Finally, for both clients and servers, \textit{Op(i).a}\(_k\), \(1 \leq k \leq n\), denotes the value of parameter \textit{a}\(_k\), and \textit{Op(i).returns} denotes the returned value.

Using the events defined above it is possible to model the different ways in which an operation can be invoked (e.g., synchronous, deferred synchronous, etc). This point is further discussed in Section 3.5.

As an example of a \textsc{corba} \textit{entity} class, let us consider class \textit{Terminal\_Obj}, which satisfies \textit{interface} class \textit{TerminalInterface}. Thus, \textit{Terminal\_Obj} inherits from \textit{TerminalInterface} and contains the axioms defining the semantics of operations \textit{getString} and \textit{getInt}.

\textbf{CORBA Entity Class} Terminal\_Obj  
\textbf{inherit} TerminalInterface  
\textbf{temporal domain} real  
\textbf{axioms}

\begin{itemize}
  \item \textbf{vars}  
    \begin{itemize}
      \item \textit{x} : \textit{string}
      \item \textit{i} : \textit{natural}
    \end{itemize}
  \item ( \textit{getString(i).start}^{10}  
    \land \textit{getString(i).maxLength} \leq 0 )  
    \rightarrow \text{WithinF} ( \textit{getString(i).negativeOrZeroPar.raise},  
    \textit{TgetString} ) \quad \text{-- If operation \textit{getString} is invoked}
  \item ( \textit{getString(i).start}  
    \land \textit{getString(i).maxLength} > 0 )  
    \rightarrow \text{WithinF} ( \exists \textit{x} ( \textit{getString(i).end\_success}  
    \land \textit{getString(i).returns} = \textit{x}  
    \land \textit{length(x)} \leq \textit{getString(i).maxLength}), \text{-- of the desired maximum length}
    \textit{TgetString} ) \quad \text{-- within \textit{TgetString} time units}
\end{itemize}

\[ ... \text{similar axioms for operation \textit{getInt} \ldots } \]

\textbf{end} Terminal\_Obj

---

\(^8\) In fact, when the \textit{POA} of the server exporting \textit{Op} receives invocation \textit{i}.
\(^9\) \textit{i} is an identifier that distinguishes invocations of the same operation that is \textit{Op(i)} and \textit{Op(j)}, \text{\textit{j} \neq \textit{i}}\, denote different invocations of \textit{Op}. Notice that we do not assume that invocations of \textit{Op} are ordered (i.e., \textit{Op(i).invoke} need not occur before \textit{Op(i+1).invoke}). If such order is needed it is necessary to add explicit axioms.
\(^{10}\) In \textit{TRIO}, free occurrences of variables are implicitly assumed to be universally quantified.
where $T_{getString}$ is a constant that bounds the response time of the operation.

As a second example consider the CORBA Entity class $GUI\_Client$ representing a Graphical User Interface (GUI) that can access two terminals, each of which supports interface $TerminalInterface$, to read either a string or an integer. When the data are collected from the terminals, they are first either concatenated (strings) or added together (integers) and then displayed. Finally, the user is notified of any error occurring because of operations invocation.

Therefore, CORBA Entity $GUI\_Client\_class$ is a client class that uses $TerminalInterface$ from which it imports operations $getInt$ and $getString$. Moreover, it introduces events $input$ and $error$ to model user interaction, state $display$ to model the output of the computation and two values, $firstTerminal$ and $secondTerminal$ to model the reference to the two terminals from which data are collected.

Notice that $firstTerminal$ and $secondTerminal$ must be initialized with the correct references before invocation of $getInt$ or $getString$ can occur. Such initialization, which can be performed using for instance the CORBA Naming Service, is not modeled in this example.

```plaintext
CORBA Entity Class GUI_Client_class
used interfaces TerminalInterface
visible input, error, display
temporal domain real
type
  PossibleOperations = {concat_string, sum_int};

TD items
  value firstTerminal : OID;
  value secondTerminal : OID;
  event input(PossibleOperations);
  event error(string);
  state display(string ∪ integer);

axioms
  vars
    i, j : natural;
    x, y : integer

  input(sum_int) → ∃i ( getInt(i).invoke ∧getInt(i).receiverID = firstTerminal )
    -- If the GUI receives a request to read integers
    -- then it invokes operation $getInt$
    -- on the first terminal

  ( getInt(i).reply ∧ getInt(i).receiverID = firstTerminal ) → ∃j ( getInt(j).invoke ∧
    getInt(j).receiverID = secondTerminal )
    -- If the GUI receives an answer
    -- from the first terminal

  getInt(i).reply ∧ getInt(i).returns = x ∧
  getInt(i).receiverID = secondTerminal ∧
  LastTime ( getInt(j).reply ∧
    getInt(j).receiverID = firstTerminal, t )
  ∧ Past(getInt(j).returns = y, t)
  → NowOn(display(x + y))
    -- If the GUI receives the value x
    -- from the second terminal and
    -- the last time it received a value from
    -- the first terminal this value was y,
    -- then it displays the sum of x and y

  [...] similar axioms for concatenating Strings ...

end GUI_Client_class
```

Instances of TC CORBA Entity classes can be either single-threaded or multi-threaded. In the default case, a CORBA Entity models a single-threaded server/client and therefore
cannot execute operations in parallel. To model multi-threaded servers/clients it is necessary to add the keyword multithreaded in the header of the class. For example, a multi-threaded server that satisfies interface TerminalInterface, is modeled by the following CORBA Entity class:

```
CORBA Entity Class parallel_Terminal_Obj multithreaded
 inherit TerminalInterface

[ ... rest of the class is omitted ... ]

end parallel_Terminal_Obj
```

A discussion of the formal semantics of keyword multithreaded is given in Section 3.5.

### 3.3 The NonCORBA Entity meta-class

NonCORBA Entity classes are used to model entities that correspond to neither servers nor clients. For example, a NonCORBA Entity class can be used to model some physical device such as a sensor not connected to an ORB, or possibly a human operator.

The syntax and the properties of NonCORBA Entity classes correspond to those of plain TRIO classes. Thus, NonCORBA Entity classes can contain, and/or inherit only from other NonCORBA Entity classes. For example, NonCORBA Entity class User_class models the behavior of a human operator requesting integer addition or string concatenation.

```
NonCORBA Entity Class User_class
 visible input, error
 temporal domain real
 type
 PossibleStates = {asleep, awake};
 PossibleOperations = {concat_string, sum_int};
 TD Items
 event input(PossibleOperations);
 event error(string);
 state state(PossibleStates);
 axioms
 [ ... axioms describing the behavior of the operator ... ]
 end User_class
```

### 3.4 The Environment meta-class

An Environment class is similar to a NonCORBA Entity class, except that it can include classes of any type. Environment classes are meant to define how the other classes composing a system interact, by describing their dependencies.

For example, let us consider a system in which several human operators (class User_class) use a graphical user interface (class GUI_client_class) to interact with two terminals (class Terminal_Obj).

The Environment class GUI_System, modeling the above-described system, is a structured class that introduces four modules: TerminalOne, TerminalTwo, GUI_client and users. TerminalOne and TerminalTwo are instances of Terminal_Obj and therefore satisfy TerminalInterface. GUI_client, which invokes getInt and getString, is an instance of GUI_client_class. Finally users is an array of instances of User_class and models N users interacting with GUI_client.

Figure 3 shows the corresponding graphical representation. Most of the symbols used are the same symbols as in TRIO. However, TC extends TRIO graphical representation by
introducing some CORBA-specific elements. For example, operations are depicted as arrows pointing from the server to the client, and interfaces are represented as boxes overlapping the servers that satisfy them.

![Diagram](image)

**Figure 3: example of graphical representation of TC classes**

In TRIO, connections define how the different modules of a structured class are linked together by stating logical equivalence among exported (visible) items. In TC, instead, some classes can export also operations. Thus, besides the usual TRIO connections between items, TC introduces the relation `bind` between operations, which is meant for modeling the behavior of the ORB. The semantics of bind is given through (pre-defined) axioms (see Section 3.5).

Visible items with the same name (events `input` and `error`) of `GUI_client` and `users` are connected together, while operations `getInt` and `getString` exported from `TerminalOne` and `TerminalTwo` are bound with the same operations imported by `GUI_client`.

```
Environment Class GUI_System
  temporal domain real
  modules
    TerminalOne : Terminal_Obj
    TerminalTwo : Terminal_Obj
    GUI_client : GUI_client_class;
    users : array [1..N] of User_class;
  connections
    (bind TerminalOne.getInt, GUI_client.getInt)
    (bind TerminalTwo.getInt, GUI_client.getInt)
    (bind TerminalOne.getString, GUI_client.getString)
    (bind TerminalTwo.getString, GUI_client.getString)
    (connect GUI_client.error, users.error)
    (connect GUI_client.input, users.input)
  axioms
    [ ... axioms describing properties of the overall system ... ]
end GUI_System
```

Notice that a TC specification always includes at least one `Environment` class, which models the system as a whole. However, since a system can be composed of different subsystems a TC architecture can comprise more than one `Environment` class in order to achieve a good modularization.

To summarize, Figure 4 shows the inherit from/contain relationships among instances of meta-classes. For example, a `CORBA Entity` class can contain only `NonCORBA Entity` classes and can inherit from `Interface`, `NonCORBA Entity` and `CORBA Entity` classes.
3.5 TC Computational Model

This section provides an overview of the semantic aspects of ORB Entity classes. The presented semantics captures the behavior of some important elements in the OMA as informally specified by the OMG. Without aiming at being exhaustive, in what follows the semantics of some aspects of operation invocation is discussed first. Then, the semantics of threads is taken into account and, finally, the semantics of the ORB is presented.

3.5.1 Semantics of operation invocations

In CORBA a client can invoke an operation in the following ways: synchronous, deferred synchronous (using the Dynamic Invocation Interface) or asynchronous\(^{11}\) (through an Asynchronous Method Invocation, as defined in the CORBA Messaging specification).

In Section 3.2 the events associated with ORB Entity classes in order to model operation invocation were introduced; next, the main axioms defining the semantics of operation invocation are presented.

The first axiom states that an operation \( Op \) can successfully reply to a client only if it has been previously invoked. Thus, the following axiom holds for every client class that imports \( Op \):

\[
\text{Op(i).reply} \rightarrow \text{SomP(Op(i).invoke)} \quad \text{[genC1]}
\]

A similar axiom exists for classes exporting \( Op \):

\[
\text{Op(i).end_success} \rightarrow \text{SomP(Op(i).start)} \quad \text{[genS1]}
\]

To model deferred synchronous operations, every client class importing an operation can invoke the built-in operations send, poll and get_response, which model the corresponding operations of the Dynamic Invocation Interface (DII) for deferred synchronous invocations. For example, \( \text{send(j)(Op(i)).invoke} \) represents the deferred synchronous invocation of operation \( Op \), which is achieved with invocation \( j \) of operation send.

From the client’s point of view, send, poll and get_response are invoked as synchronous operations. Thus, the following axiom states that operation send eventually ends (either successfully or with an exception), within \( T_{ORB} \) time units, where \( T_{ORB} \) is an ORB-dependent constant.

\(^{11}\) Asynchronous Method Invocations are a recent addition to the CORBA standard, which have not yet been introduced in any CORBA implementation. Therefore they are not considered in this paper.
\[\text{send}(j)(\text{Op}(i)).\text{invoke} \quad \text{[defsynC1]}\]
\[\rightarrow \text{WithinF} (\text{send}(j)(\text{Op}(i)).\text{reply} \oplus \exists \text{Exc}^{12} (\text{send}(j)(\text{Op}(i)).\text{Exc}.\text{received}, \text{T}_{\text{ORB}}))^{13}\]

Notice that \text{T}_{\text{ORB}} is an abstraction that hides all the internal aspects of the ORB, such as the transport layer, the operating systems timeouts etc.

The difference between a deferred synchronous invocation and a synchronous one is that the former is modeled as a synchronous invocation to the DII of the ORB, which returns almost immediately, while the latter is modeled as an invocation to a server, which is not guaranteed to return (e.g., if the server crashes).

Since the results of a deferred synchronous invocation can be retrieved only after the operation is invoked, client classes contain, for each imported operation \text{Op}, the following axiom:
\[\text{get}_r\text{esponse}(j)(\text{Op}(i)).\text{reply} \rightarrow \exists k (\text{SomP}(\text{send}(k)(\text{Op}(i)).\text{reply}) \quad \text{[defsyncC2]}\]

Finally, \text{get}_r\text{esponse} is a blocking operation, which returns control to the invoking client \text{C} only after the corresponding remote operation terminated on the server \text{S}. As a consequence, an \text{Environment} class containing \text{C} and \text{S} includes the following axiom:
\[\text{C.get}_r\text{esponse}(j)(\text{Op}(i)).\text{reply} \rightarrow \text{SomP}(\text{S.Op}(i).\text{end}_s\text{uccess} \oplus \exists \text{Exc} (\text{S.Op}(i).\text{Exc}.\text{raise}) \quad \text{[defsyncC3]}\]

### 3.5.2 Thread semantics

Instances of \text{CORBA Entity} classes are associated with one or more threads, each of which can be in one of the following states: \text{idle}, \text{busy} or \text{blocked}. This is formalized by introducing predicate \text{thread}(k), \text{k} \geq 0, which states that \text{k} threads are associated with a given instance, and predicate \text{thread}(j).\text{state}(s), s \in \{\text{idle}, \text{busy}, \text{blocked}\}, which states that thread \text{j} is in state \text{s}.

Since operation invocations (received or made by a \text{CORBA Entity} instances) are carried out by threads, predicate \text{Op}(i).\text{thread}(j) is true if and only if invocation \text{i} of operation \text{Op} is handled by thread \text{j}.

Variable \text{k} in \text{thread}(k) equals 1 for a single-threaded server/client, and thus all operations are associated with the same thread. For a multi-threaded server/client with a threadpool of \text{n} threads, variable \text{k} can range from 1 to \text{n}, while \text{k} is unbounded if no number of threads is specified.

A client thread that synchronously invokes an operation remains \text{blocked} until the operation returns (successfully or not):
\[\text{thread}(k).\text{state}(\text{blocked}) \leftrightarrow \exists \text{Op}, i (\text{Op}(i).\text{thread}(k) \wedge \text{Since}(\neg \text{Op}(i).\text{reply} \wedge \neg \exists \text{Exc} (\text{Op}(i).\text{Exc}.\text{received}), \text{Op}(i).\text{invoke})) \quad \text{[threadC1]}\]

where \text{Op} ranges over all imported operations of the client class (including \text{send}, \text{poll} and \text{get}_r\text{esponse}).

When a thread is blocked, it cannot invoke operations:
\[\text{thread}(k).\text{state}(\text{blocked}) \rightarrow \neg \exists \text{Op}, i (\text{Op}(i).\text{thread}(k) \wedge \text{Op}(i).\text{invoke}) \quad \text{[threadC2]}\]

\[^{12} \exists \text{Exc} (\text{Op}(i).\text{Exc}.\text{received}) \text{ is a shortcut for } \text{Op}(i).\text{Exc}1.\text{received} \lor \ldots \lor \text{Op}(i).\text{Exc}N.\text{received}, \text{ where Exc}1 \ldots \text{Exc}N \text{ are the exceptions that can be raised when invoking Op.}\]
\[^{13} \text{Symbol } \oplus \text{ stands for the “exclusive or”.}\]
Therefore, a single-threaded client remains blocked while waiting for a synchronous invocation to terminate.

On the server side, when an invocation is received and all available threads are busy or blocked the incoming request is queued and will be served only when one thread becomes idle. The axioms defining the queuing of incoming invocations are omitted for sake of brevity.

### 3.5.3 Modeling the ORB

When an operation is invoked by a client, the ORB dispatches the request to the server. In general, a delay (because of both the ORB and the underlying network) occurs between the moment when the client sends the request (through the ORB), and the moment when the server receives the invocation.

Let $Op$ be an operation invoked by client $C$ on server $S$ (where $C$ and $S$ are instances of some CORBA Entity classes). The following axiom states that when $C$ invokes $Op$, either the invocation is eventually dispatched to $S$, or a standard exception $COMM\_FAILURE$ is raised by the ORB:

\[
C.\text{Op}(i).\text{invoke} \land C.\text{Op}(i).\text{receiverID}(S._\text{id}) \quad \rightarrow \quad \begin{cases} 
\text{SomF}(S.\text{Op}(i).\text{inv\_received} \land S.\text{Op}(i).\text{callerID}(C._\text{id})) \\
\oplus \text{SomF}(C.\text{Op}(i).\text{COMM\_FAILURE}.\text{received}) 
\end{cases}
\]

where $C.\text{Op}(i).\text{receiverID}(S._\text{id})$ is true if and only if $Op$ is invoked on server $S$. Conversely, axiom [ORB2] states that if $S$ receives an invocation from client $C$ ($S.\text{Op}(i).\text{callerID}(C._\text{id})$ is true), the invocation was previously made by $C$:

\[
S.\text{Op}(i).\text{inv\_received} \land S.\text{Op}(i).\text{callerID}(C._\text{id}) \quad \rightarrow \quad \text{SomP}(C.\text{Op}(i).\text{invoke} \land C.\text{Op}(i).\text{receiverID}(S._\text{id}))
\]

Finally, axiom [ORB3] states that when $C$ receives a reply concerning invocation $i$ of $Op$, the operation successfully terminated on $C$.

\[
C.\text{Op}(i).\text{reply} \land C.\text{Op}(i).\text{receiverID}(S._\text{id}) \quad \rightarrow \quad \text{SomP}(S.\text{Op}(i).\text{end\_success} \land S.\text{Op}(i).\text{callerID}(C._\text{id}))
\]

These (and other) axioms define the meaning of bindings between operations and therefore belong to the Environment class containing $C$ and $S$.

For each binding between operations defined in an Environment class, an instance of the previous axioms is automatically generated. For example, class $GUI\_System$ contains, among others, the following axiom referring to operation $\text{getInt}$ between $TerminalOne$ and $GUI\_client$:

\[
(\quad GUI\_client.\text{getInt}(i).\text{invoke} \quad \land 
\quad GUI\_client.\text{getInt}(i).\text{receiverID}(TerminalOne._\text{id})) \\
\rightarrow \quad \begin{cases} 
\text{SomF}(\quad TerminalOne.\text{getInt}(i).\text{inv\_received} \\
\quad \land TerminalOne.\text{getInt}(i).\text{callerID}(GUI\_client._\text{id})) \\
\oplus \text{SomF}(GUI\_client.\text{getInt}(i).\text{COMM\_FAILURE}.\text{received})) 
\end{cases}
\]

[GUI\_ORB2] and [GUI\_ORB3] are generated similarly.

Notice that there is no axiom that guarantees that after the operation ends on the server, the answer is sent back to the client, since there could be a communication failure that prevents this to occur.

The delay between $C.\text{Op}(i).\text{invoke}$ and $S.\text{Op}(i).\text{inv\_received}$ (and between $S.\text{Op}(i).\text{end\_success}$ and $C.\text{Op}(i).\text{reply}$) remains unbounded since we do not deal with real-time constraints (and
neither the ORB, nor the network give guarantees about maximum latencies); however, this can be further refined by establishing an upper bound for such a delay (see [Marotta et al. 2001] for an analysis and formalization of real-time CORBA and [Pradovera 2001] for a first qualitative and experimental assessment thereof).

Moreover, there could be a further delay before the server actually starts processing the invocation. For example, if a single-threaded server is busy because of another invocation (from another client) it queues the incoming request, and will process it whenever possible.

The axioms modeling the behavior of servers between the moment when they receive an invocation and the moment when the operation starts being processed are not shown for the sake of brevity.

3.6 Stereotypes

Stereotypes can be used to extend the semantics of TC elements. They can be attached to any element in a TC diagram, once they are defined. Defining a stereotype means to state, using a suitable syntax, the new axioms that must be associated with the marked element along with the existing axioms that do not hold anymore, if any.

For example, let us consider CORBA oneway operations, which must obey some signature constraints (they only have in parameters, do not return any value, they cannot raise user-defined exceptions), and adopt a best-effort invocation semantics. In order to define such operations it is possible to introduce the stereotype «oneway», which can be associated with exported operations. The «oneway» stereotype does not introduce any new axioms, but requires that axiom [ORB1] does not hold, since the invocation is not guaranteed to be dispatched to the server.

As a second example, let us assume that one needs to model reliable servers, that is fault tolerant servers exporting operations guaranteed to return (provided that they do not enter an infinite loop) with either a successful result or an exception\(^{14}\). Notice that the axioms of Section 3.5 do not guarantee that an operation returns after being invoked since the server could crash before the operation has ended. In order to define such reliable behavior it is possible to introduce the stereotype «reliable» that can be associated with any server class. The «reliable» stereotype extends the properties of a CORBA Entity class, as defined by the following declaration, which introduces a new axiom for each exported operation Op:

```plaintext
stereotype reliable applies to CORBA Entity
foreach Op in Exported
  add
    vars:
      i : natural;
      Reliable:
        Op(i).start → SomF (Op(i).end_success ⊕ ∃Exc (Op(i).Exc.raise))
end
```

Stereotype «reliable» (and the associated semantics described by axiom [Reliable]) is applied to servers, and concerns only the exported operations. Notice that, from the client’s viewpoint, reliability requires both the server and the underlying ORB to be reliable.

As a final example, consider the CORBA Event Service [OMG 2001b], which provides a way for exchanging (CORBA) events among objects interacting through the ORB. Objects sending

\(^{14}\) One way of building reliable servers is by means of the Object Group Service [OpenDREAMS-II 1998b] developed in OpenDREAMS-II, which provides a way for managing a set of replicated servers. Moreover, the CORBA standard includes the specification of a Fault-Tolerant CORBA.
events are referred to as suppliers, while objects receiving events are referred to as consumers. An event can be viewed as a chunk of information concerning something that happened in the supplier (e.g., an attribute changing value) that is of some interest to one or more consumers. The Event Service allows one to decouple suppliers from consumers by introducing event channels. Thus, rather than having the supplier invoking an operation of the consumer to notify that the event occurred (or having the consumer invoking an operation of the supplier to ask whether the event occurred), the Event Service provides event channels to dispatch events from suppliers to consumers. OMG specifies the following models for initiating the event dispatching:

- the supplier initiates the event dispatching by invoking an operation on the event channel (supplier push model), or
- the event channel initiates the event dispatching by invoking an operation of the supplier to check whether an event occurred (supplier pull model).

In a similar way, it specifies a push model and a pull model for the consumer:

- the event channel notifies the consumer by invoking one of its operation (consumer push model), or
- the consumer queries the event channel to check whether any event is available (consumer pull model).

In order to formalize the Event Service, TC introduces four different stereotypes representing the different combinations of push and pull models for suppliers and consumers. For example, if operation \( Op \) exported by server \( S \) and imported by client \( C \) is marked as «eventPush», its invocation corresponds to an event dispatching made using the push model for both the supplier and the consumer. Conversely, if \( Op \) is marked as «eventPull», the event is dispatched using the pull model for both the supplier and the consumer. Notice that, in the former case, \( S \) is the consumer and \( C \) is the supplier while in the latter case \( S \) is the supplier and \( C \) the consumer. In both cases, event dispatching occurs asynchronously, that is, termination of operation \( Op \) on the client is decoupled from termination on the server.

Each of the stereotypes mentioned above formalizes the behavior of the Event Service according to the different models of event dispatching.

For example, an Environment class \( E \), containing both \( S \) and \( C \), in which \( Op \) is marked as «eventPush» will not contain axiom [ORB3] for \( Op \), which relates the termination of \( Op \) for the client (i.e., \( C.Op(i).reply \)) with the actual termination of \( Op \) on the server (i.e., \( S.Op(i).end\_success \)). In fact, \( C \) invoking \( Op \) represents

1. the supplier (\( C \)) invoking an operation on the event channel to push the event into the event channel, and
2. the event channel invoking \( Op \) to dispatch the event to the consumer (\( S \)).

Since the termination of event dispatching does not occur before termination of event pushing, axiom ORB3 does not hold for \( Op \). Instead, other axioms, not discussed here for the sake of brevity, are added to formalize the behavior of the Event Channel. The interested reader can refer to [Rossi 2002] for a more detailed definition of stereotype «eventPush».

4 THE TC METHODOLOGY

High level design essentially consists of identifying the classes composing the system whose instances provide and use services by interacting through the ORB. The design of the high level architecture of an application is influenced by the outcome of the specification phase: a good design is very often the natural consequence of a good specification. On the other hand,
during the design it may be necessary to reconsider some of the requirements stated in the specification. As a result, the design phase should not be considered as a self-contained process. This paper, however, does not discuss the issue of writing, analyzing and managing specifications; the interested reader can refer to [Ciapessoni et al. 1999] for an industrial experience on specifying system requirements using TRIO.

This section presents a methodology, named TRIO/CORBA Methodology (TCM) to design the high level architecture of a CORBA-based system starting from a TRIO specification. The goal of this methodology (as well as of other methodologies) is to show the conceptual path that one has to follow in order to design the architecture of a system. Along the path, designers have to make design choices that influence the resulting architecture. Although we believe that no methodology should be considered as a “silver bullet”, we think that a categorization of the different design choices along with some guidelines on how to tackle such choices can indeed be useful to designers. Therefore, a good methodology should:

1. identify the different conceptual problems that designers have to face;
2. suggest the best order in which such problems should be tackled;
3. provide guidelines and suggestions on how to solve the most critical problems that designers will face.

TCM allows designers to smoothly move from the specification towards the high level design in a step-wise fashion. At each step, a different aspect is taken into account so that the complexity of the whole design is kept under control. Moreover, at each step a “design document” is produced using TC, in order to keep track of the different choices made.

In the presentation of TCM, we point out which steps are straightforward and therefore can be automated, and which ones cannot be performed without a creative effort from the designer. In what follows the steps are presented as if they were meant to be executed sequentially. However it is useful to remind that they are not completely independent and that, in practice, they follow an incremental approach [Mills et al. 1987]. Moreover, mutual feedbacks among the various phases and sub-phases are unavoidable according to the philosophy of the spiral approach [Boehm 1988].

The methodology is mainly structured into five major steps:

- identification of data flows between the specification classes;
- identification of operations and attributes;
- identification of objects;
- identification of interfaces and semantics of operations and attributes;
- identification of services and non-architecture-impacting frameworks.

The methodology is illustrated by means of an example based on a Maintenance System developed by ENEL, the Italian energy agency, within the ESPRIT Project OpenDREAMS-II [OpenDREAMS-II 1998a]. In what follows the Maintenance System is introduced and a short discussion of its TRIO specification is provided. Then, the different steps of the methodology are presented in detail.

Notice that, when dealing with industrial projects, one has often to face legacy aspects of (already deployed) systems, which can, in some cases, limit the number of choices that can be made during the different phases of application development.
4.1 The ENEL Maintenance System

The goal of the Maintenance System (MS) is to monitor the activity of field devices (sensors, actuators, etc.) installed in a power plant, in order to quickly detect possible failures and malfunctions.

The core of the system is the Instrumentation Maintenance System (IMS), which is in charge of collecting and validating data (i.e., measurements) coming from the field devices. Whenever the validation process detects an anomaly in the behavior of such devices, IMS sends an alarm to Alarm Manager (AM), which in turn notifies a human operator by means of a Human-Machine Interface (HMI).

Figure 5 shows the main components of the application and their mutual interactions.

![Figure 5: The Maintenance System](image)

Notice that IMS does not communicate directly with the field devices: all the data collected by these devices are stored into a database named Global Plant Database (GPDB). Thus, IMS queries GPDB to obtain the desired data. Using the same communication mechanism IMS can also send commands to the devices or can make a device perform a self-test to verify its correct functioning. However, before sending a command to a device, IMS must get from the Control System (CS) the rights to access such device. After having completed the desired operations, IMS notifies CS, which, in turn, releases the device.

For the sake of simplicity, this paper does not take into account the interaction between the user and HMI.

The final, complete implementation of the MS application consisted of 18 classes and 28 IDL interfaces. The actual number of objects depends on the number of connected devices (in the first field tests this number was about 60). Communication is mostly performed through synchronous operations.

The only parts of the application that do not rely on CORBA for communication are the field devices. ENEL imposed as a design decision the use of field-bus [IEC] for managing the interaction among devices and the GPDB. This decision was strongly based on the state of the art at the time of the project.

4.2 The TRIO Specification

The specification of the Maintenance System consists of a single structured TRIO class (MaintenanceSystem), modeling both the system as a whole and the environment in which the

---

15 The number of IDL interfaces includes those defining static information such as Operators roles, Types of values etc, which are not presented in the paper.

16 A field-bus is the typical SCS digital channel used to connect sensors and other equipment to computers. The choice of a field-bus was done in order to be compliant with the current ENEL internal standards.
system has to operate. Such class, in turn, comprises modules modeling the different components of the system.

Most of the identified classes come directly from the informal specification of the system. Thus, the specification contains classes modeling the control system, the instrumentation maintenance system, and so on. The only part of the specification not directly coming from the informal specification is the one that describes the way in which IMS, Alarm Manager and HMI interact. In fact, the most natural way to specify such interaction would be to introduce a class (AlarmDispatcher) in charge of dispatching the alarms detected by IMS to HMI, as shown in Figure 6.

![Figure 6: The alarm-dispatching system](image)

Instead, specifiers decided to manage alarms by using the Anomalies Detection Module (ADM), which is a framework developed in OpenDREAMS-II whose aim is to notify an operator through a human-machine interface of the occurrences of anomalies in some other module [Capobianchi et al. 1999, OpenDREAMS-II 1998c]. However, in order to use ADM, each source of alarm (i.e., sensors presenting anomalies) must be associated with a different object within ADM. Such objects are in charge of raising alarms whenever the corresponding entities enter an abnormal state. Therefore, the specification of the alarm management part must be structured as reported in Figure 7.

![Figure 7: The alarm-dispatching system using ADM](image)

Although the goal of the specification is to describe what the system must do without describing how it can be implemented, in many industrial environments separation between application requirements and implementation choices is not as sharp, whether this is useful or dangerous [Jackson 1995]; This is also due to the legacy aspects that we expressed at the beginning of this section. An example of this overlapping is represented by the use of architecture-impacting frameworks, such as ADM, that contain in their very definition architecture-shaping concepts, so that their use should be carefully considered since the specification phase.

Figure 8 shows the graphical representation of the TRIO classes composing the specification for the part of the system taken into account. Notice that the part affected by the introduction of ADM has been enclosed in a box. This can be viewed as a reminder for the following steps in order to make the design choices specifically tailored towards the use of such framework.

Figure 8 shows also some of the TRIO items used to specify the behavior of classes and the way in which they are connected. For example, item test_request is an event that is true when IMS asks a device, via GPDB, to perform a self-test, while access_avail is a non-visible state representing whether or not IMS has acquired the access rights from CS. The Validation module contained in IMS identifies the part devoted to validating the measurements coming from the devices.
The behavior of the system is expressed by means of axioms providing semantics to the different classes. For instance, the following axioms of class IMS state that:

- if a test \((\text{test}_\text{cmd})\) is started \((\text{test}_\text{request})\) or any other command \((\text{dev}_\text{cmd})\) is sent \((\text{command}_\text{send})\) to a device \(MC\), then IMS must have already acquired the access rights from CS \((\text{access}_\text{avail})\);
  \[
  \left(\text{test}_\text{request}(i, MC, \text{test}_\text{cmd}) \lor \text{command}_\text{send}(i, MC, \text{dev}_\text{cmd})\right) \rightarrow \text{access}_\text{avail}
  \]

- if the testing activity \((\text{test}_\text{cmd})\) on a device ends \((\text{test}_\text{end})\) then it was previously started \((\text{test}_\text{request})\);
  \[
  \text{test}_\text{end}(i, MC) \rightarrow \exists \text{test}_\text{cmd} \left(\operatorname{SomP}(\text{test}_\text{request}(i, MC, \text{test}_\text{cmd}))\right)
  \]

- If (and only if) IMS requests CS the rights to access the devices \((\text{access}_\text{request})\) and after sixty seconds no answer \((\text{access}_\text{granted} \lor \text{access}_\text{denied})\) has been received from CS, then IMS issues a request to abort the process \((\text{abort}_\text{request})\).
  \[
  \exists j \left(\text{abort}_\text{request}(j) \leftrightarrow \exists i \left(\operatorname{LastTime}(\text{access}_\text{request}(i), 60) \land \operatorname{Lasted}(\neg(\text{access}_\text{granted}(i) \lor \text{access}_\text{denied}(i)), 60)\right)\right)
  \]

In turn, the following axiom of CS states that when IMS issues an \(\text{abort}_\text{request}\), previous unanswered access requests are aborted, that is neither \(\text{access}_\text{granted}\), nor \(\text{access}_\text{denied}\) will be issued.

\[
\left(\text{abort}_\text{request}(j) \lor \neg(\text{access}_\text{granted}(i) \lor \text{access}_\text{denied}(i))\right) \rightarrow \neg\operatorname{Som}(\text{access}_\text{granted}(i) \lor \text{access}_\text{denied}(i))
\]

Furthermore, the following axioms of class GPDB state that:
If GPDB sends to IMS the status \((\text{dev\_s})\), the operating mode \((\text{om})\) and the access permission \((\text{ac\_p})\) concerning a device \(\text{MC}\) (\(\text{cyclic\_acq}\)) by means of \(\text{chan\_status}\), then it has acquired such data from MC by means of \(\text{status}\) in the last \(T_{\text{max}}\) time units:\(^{17}\)

\[
\begin{align*}
\text{cyclic\_acq}(i, \text{MC}) & \quad \rightarrow \exists \text{dev\_s}, \text{om, ac\_p} \quad ( \text{chan\_status}(\text{MC, dev\_s, om, ac\_p}) \\
& \quad \wedge \text{WithinP} (\text{status}(\text{MC, dev\_s, om, ac\_p}), T_{\text{max}}))
\end{align*}
\]

If GPDB receives the request to initiate a test \((\text{test\_cmd})\) on device \(\text{MC}\), the test will eventually end.

\[
\text{test\_request}(i, \text{MC, test\_cmd}) \rightarrow \text{SomF} (\text{test\_end}(i, \text{MC})) \quad [\text{ax6}]
\]

An (early) complete TRIO specification of the MS is given in [Pradella 2000].

### 4.3 From the Specification to the Design

In what follows the design methodology is illustrated by applying it to the Maintenance System.

#### 4.3.1 Step 1: Data Flows

The first step aims at identifying explicit information exchanges among the classes identified in the specification. These exchanges are called \textit{data flows} and are a first step to move from the concept of sharing logical items (predicates, functions, etc) – typical of TRIO classes – towards the concept of exported operations – typical of CORBA.

A data flow can be viewed as a complex merge of TRIO items and is either \textit{unidirectional} or \textit{bi-directional} depending on the direction in which information flows. The decision on how to regroup TRIO items is taken by analyzing the behavior (i.e., the axioms) of the classes to which the items belong. For example, let us consider axioms [\text{ax2}] and [\text{ax6}]: combined together, they state that \(\text{test\_end}\) is true if and only if \(\text{test\_request}\) was true sometime before. Furthermore, the following axiom (from class GPDB)

\[
\text{test\_end}(i, \text{MC}) \\
\rightarrow \exists \text{mval, vi, timetag} \quad ( \text{measure\_info}(\text{MC, mID, mval, vi, timetag}) \\
& \quad \wedge \text{WithinP} (\text{measure}(\text{MC, mID, mval, vi, timetag}), T_{\text{max}}))
\]

states that, when a test ends, \(\text{measure\_info}\) “carries” the value that was exchanged at most \(T_{\text{max}}\) time units ago between GPDB and MeasuringChannel. Since similar axioms relate \(\text{test\_end}\) with \(\text{chan\_status}\) and \(\text{chan\_detailed\_status}\), we group all of them into a bi-directional data flow named \(\text{test}\). In fact, \(\text{test\_end}\) denotes the end of a test whose beginning is represented by \(\text{test\_request}\), while \(\text{measure\_info, chan\_status}\) and \(\text{chan\_detailed\_status}\) describe the results of the test.

The semantics of \(\text{abort\_request}\), instead, suggests that there is a unidirectional flow from IMS to CS. In fact, if IMS issues an \(\text{abort\_request}\) ([\text{ax3}]) then CS does not send back any answer (it actually stops ongoing processing of any pending \(\text{access\_request}\), as stated by [\text{ax4}]).

Both examples show that whenever some items are related (by means of axioms) to describe a cause-effect relationship they become candidates to be transformed into a dataflow. Thus, deciding whether to perform such transformation is a typical design choice, since data flows are meant for modeling ORB-based communication.

---

\(^{17}\) \(T_{\text{max}}\) is a system-dependent constant representing the maximum delay between the instant when data are collected from the devices and the instant when they are sent to IMS.
For example, items *measure*, *status* and *detailed_status*, connecting classes GPDB and MeasuringChannels, are not grouped into a data flow. In fact, they represent the information exchanged between the devices and GPDB, and the design choice made is to use a field-bus to make the two entities communicate, instead of an ORB. Therefore, they are not grouped into a data flow and their representation remains as it was in the specification. However, the field-bus imposes to introduce a new item (*ctrl*) connecting GPDB with the devices, representing a control signal. In fact only when *ctrl* is true, *measure*, *status* and *detailed_status* have meaningful values that can be accessed by GPDB.

Notice that, in the future, the same situation might call for a different solution such as using a real-time ORB. In such a case those same items should be grouped into a data flow representing ORB-based information exchange between GPDB and field devices.

A cause-effect relationship between events is not the only reason to group items into a data flow as the following example shows.

Suppose that the anomalies detected by IMS are categorized into light and severe. Thus, the specification would contain two events *warning* and *alarm* connecting MCStates and HMI and representing light and severe anomalies, respectively. Upon closer inspection of the semantics of these events, the designer might decide that they represent the same flow of information (some notification from MCStates to HMI), and group them in a unique data flow *alarm_send*.

In this case, grouping is not the consequence of an existing cause-effect relationship among items; rather, it depends on the fact that such items describe a similar flow of information. Thus, the designer may find that the specification introduced different items to model events that are conceptually similar, and he/she can decide to abstract from such initial view grouping them into a single data flow.

Figure 9 reports the TRIO classes of the specification of figure 8 in which some items have been replaced by data flows.

Moreover, every data flow is textually defined. For example, the definition of *test* is as follows:

![Figure 9: Data Flows representation](image-url)
Connection between IMS and GPDB

Dataflows

test (from test_request,
    to test_end,
    to chan_status,
    to chan_detailed_status,
    to measure_info);

test is then a bi-directional data flow (it contains both from and to elements), where test_request represents information flowing from IMS to GPDB, and all other items correspond to an information flow from GPDB to IMS.

4.3.2 Step 2: Clients and Servers

In the second step, every data flow is categorized as either operation or attribute. For each operation, one has to choose the class that exports it (server class) and the classes that import it (client classes). Similarly, for each attribute, one has to choose the class that declares it and the classes that access it.

Notice that the direction of a data flow connecting two classes does not determine a priori the client class and the server class. This is a typical design choice that concerns the communication style that one wants to use.

For example data flow test becomes an operation (with the same name) exported by GPDB and invoked by IMS, since test must be executed only when IMS issues a command.

In Figure 10 arrows identifying the operations are drawn from exporting classes to importing ones.

Figure 10: The Maintenance System after steps 1 and 2

GPDB exports two other operations, command (derived from flow command_send) and get_measure (derived from flow cyclic_acq), while it invokes the operation variation (derived from flow on_variation_acq) exported by IMS. Since we decided to exploit the Anomalies Detection Module, dataflow change_state has become operation set_status of that module.
4.3.3 Step 3: Objects

This step aims at identifying all CORBA Entities (i.e., servers and clients) that need to be implemented. Such identification is based on the operations/attributes introduced in the previous step.

Every class exporting/importing at least one operation (attribute) is a natural candidate to become an instance of the CORBA Entity meta-class since it will use the ORB to communicate. Notice that each CORBA Entity class has to satisfy the corresponding axioms of the specification. However, since in the previous steps TRIO items have been merged into data flows it is necessary to rewrite such axioms. This point is further discussed in Section 4.3.6.

Identifying the right CORBA Entity classes is a crucial design step in which the actual structure of the system is defined. Thus, the problem is to decide to which extent the structure of the specification should be maintained when designing the high level architecture of a system. Notice that, although TRIO supports the object-oriented paradigm, the experience has shown that very often specifiers tend to give a functional-oriented specification. This is not a bad practice per se, but may lead to a class structure that needs to be modified in order to identify the actual CORBA objects. Thus, in order to come up with a real object-oriented architecture it may be necessary to split and/or group some of the classes of the specification.

For example, classes IMS and GPDB are candidate to become CORBA Entity classes since they both export at least one operation. However, after having analyzed the axioms of module GPDB, we decided to divide it into two parts, named Gateway and DataRep, as shown in Figure 11. Looking at the behavior of class GPDB we notice that the axioms referring to operations test and command are independent from those referring to get_measure and variation. In fact, Gateway acts as a gateway for sending commands while DataRep acts as the actual database, storing all the measurements collected by the devices. Therefore, IMS, Gateway and DataRep are CORBA Entity classes.

![Figure 11: The CORBA Entity classes](image)

The rationale underlying the splitting of GPDB into Gateway and DataRep is that GPDB can be decomposed into two different parts playing different roles. This kind of decomposition is a design choice concerning the actual structure of the overall system, which in general may be different from the structure of the specification.
Thus, the decision of splitting a class is taken by analyzing its behavior (i.e., the axioms). If the behavior of part of a class is independent from the rest of it then the class is a good candidate for splitting. A class is composed of independent parts if it is possible to identify a subset of its axioms that do not refer to the items occurring in the remaining axioms (and vice versa).

Notice that also operations test and command of Gateway are independent since the events defining command do not appear in the axioms that refer to test and vice versa. However, in this case the design choice was to keep these two operations in the same server (Gateway), since they both provide a way to issue commands to the measuring devices.

Another reason for modifying the class structure of the specification is the introduction of an architecture-impacting framework. However, this was not our case since the framework has been taken into account at specification-time.

In conclusion, restructuring the classes composing a system is a major design choice that belongs to the designer who acts on his/her knowledge of what the system does. Thus, a “mechanical analysis” of the axioms can help in identifying the classes that are candidate for splitting/grouping but, as usual, the final choice is left to the designer.

Class MeasuringChannels does not correspond to any CORBA object, since it does not interact with the rest of the application by means of CORBA operations and/or attributes (see Section 4.3.2). Therefore, MeasuringChannels becomes an instance of the NonCORBA Entity meta-class.

### 4.3.4 Step 4: Interfaces and Semantics of Operations and Attributes

All CORBA Entity classes identified in the previous step acting as server classes must be provided with the needed interfaces. This is done by introducing instances of the Interface meta-class and making the CORBA Entity classes exporting at least one operation/attribute inherit from them. In our example, six different interfaces have been introduced (one for each CORBA Entity class) as shown in figure 12.

![Figure 12: The TC diagram after step 4](image)

Notice that Interface `ADM::State` is a standard interface defined in the Anomalies Detection Module, which is satisfied by CORBA Entity MCstates.
Providing the interface to server classes, unlike other design choices, can be done in a systematic way and can be performed almost automatically. Thus, once the CORBA Entity classes (and their interfaces) have been identified the structure of the architecture is defined.

Let us consider now a server class exporting both operations and attributes. According to CORBA specifications, the default is that operation invocations have either an at-most-once semantics, if they raise an exception, or an exactly-once semantics if they return successfully, while attributes can be accessed and modified by all clients importing them.

In addition, CORBA allows operations to be declared oneway, which means that a best-effort semantics, which does not guarantee delivery of the request, is adopted; also, attributes can be declared read-only, which means that clients cannot modify them.

Thus, during this step the designer can add the TC stereotypes «oneway» and «readonly» to identify oneway operations and read-only attributes, respectively.

For example, we decided to use for the access_yield operation a best-effort semantics, so we marked it «oneway» as shown in Figure 12.

4.3.5 Step 5: Services and Frameworks

CORBA Services and Frameworks are defined by means of a set of IDL interfaces, and thus they are modeled by means of Interface classes. Notice that translating IDL interfaces into Interface classes (and vice versa) can be done automatically, since they only differ from a syntactic point of view.

Conversely, in order to provide semantics to a CORBA Service (Framework) one has to define the underlying CORBA Entity classes satisfying the Interface classes. In other words, it is necessary to formalize in TC the semantics of the CORBA Services (Frameworks), and possibly of further independent ones, starting from the informal specification issued by the OMG (or the independent developer).

However, in order to introduce in TC a CORBA Service (Framework) one has to (automatically) translate its IDL interfaces into Interface classes without being compelled to provide also the formal semantics of such Service in term of CORBA Entity classes. Thus, there is a trade-off between the amount of effort needed to introduce existing CORBA Services (Frameworks) and the level of formality obtained, which in turn defines the level of formal correctness that one may prove. This is an example of the lightweight approach that we advocate and that is discussed in depth in a previous paper [Ciapessoni et al. 1999].

From a graphical point of view, CORBA Services are normally represented by means of stereotypes, while Frameworks are usually represented by a class structure.

When designing the Maintenance System, the CORBA services taken into account were event, transaction, query, object group and persistence; and a TC formalization has been made for some of them [OpenDREAMS-II 1998b]. Object group and persistence are used by objects while query and transaction involve operations on objects. All these services can cooperate in order to allow an object to fulfill its requirements.

For example, DataRep is a critical component and needs to be reliable to satisfy the fault tolerance requirements of the system. In the specification, data acquisition from GPDB was modeled by means of events (cyclic_acq and on_variation_acq), and the fault-tolerance requirement was implicitly stated. However, as more low-level details are included in the design, the requirement had to be made explicit, and therefore the stereotype «reliable» (see Section 3.6) was added to DataRep. Therefore, the following axiom (instance of [Reliable]) is automatically introduced:
get_measure(i).start \[get\_mReliable]\n\n→ SomF (get_measure(i).end_success \oplus \exists \text{Exc} (get_measure(i).Exc.raise))

Furthermore, let us consider operation variation invoked by DataRep to notify IMS that an abnormal variation of some measured quantity occurred. Such operation should not block DataRep while IMS processes the information since other abnormal variations may occur. This can be obtained in different ways. For instance

1. DataRep is a multithreaded client. However, this would raise the complexity of the connection with the measuring channels, or
2. DataRep invokes variation in a deferred synchronous way, or

Solution 3) is better than solution 2) if DataRep has to notify abnormal variations to objects other than IMS. In such a case solution 2) requires that DataRep invoke an operation for each object, while solution 3) requires that DataRep invoke only the CORBA Event Service.

Since in the complete system other components than IMS should be notified we decided to use the CORBA Event Service. Furthermore, dataflow on_variation_acq (see Figure 9) shows that the information flows from DataRep to IMS that is DataRep is the supplier while IMS is the consumer. The design choice of transforming on_variation_acq into operation variation, exported by IMS and imported by DataRep, leads to adopt the push model for both supplier and consumer. Thus, operation variation is marked with the stereotype «eventPush» as shown in Figure 13.

Notice that ADM relies on CORBA Event Service to deliver alarm notifications, and it defines a special type of event (AlarmEvent), which is exchanged between alarm suppliers and alarm consumers [OpenDREAMS-II 1998d]. Operation raise_alarm, dispatching alarms from MCstates to HMI, is implemented using ADM and thus, is labeled with stereotype «ODAlarmEvent».

![Figure 13: The final TC diagram](image)

Finally, the Base Process Value framework [Capobianchi et al. 1999], defined and implemented in OpenDREAMS-II, is introduced. This framework provides a way to store and manipulate the values coming from devices along with some related information such as time.
stamps and validity. It is meant for SCSs and it defines several different interfaces, one of which \((od\text{\texttt{Float}})\) is used in the example by DataRep, Gateway and IMS to exchange information about the measured values.

At the end of this step, the IDL interfaces are automatically produced from the CORBA Interface classes.

### 4.3.6 Tuning up the axioms

Once the structure of the system architecture is defined, one can express the semantics of the different classes by adapting the axioms of the specification in order to take into account all the transformations that have occurred. During this process, one must make sure that the requirements captured by the (formal) specification are still satisfied by the architecture. In what follows, we outline the transformation that should be performed on the axioms in the specification to come up with the axioms that reflect the design choices made.

For example, during steps 1 and 2 items \textit{test\_request} and \textit{test\_end} were associated with the invocation of operation \textit{test} and the moment when this operation returns, respectively. Thus, \([ax1]\) is transformed into the following axiom of class IMS in which data flows are involved (item \textit{access\_avail} remained unchanged since it does not belong to any data flow).

\[
(test(i).\text{invoke} \lor \text{command}(i).\text{invoke}) \rightarrow \text{access\_avail} \quad [ax1']
\]

Similarly, axiom \([ax2]\) would become:

\[
\text{test}(i).\text{reply} \rightarrow \text{SomP}(test(i).\text{invoke}) \quad [ax2']
\]

however, \([ax2']\) can be dropped since it is implied by the definition of synchronous operation ([\textit{genC1}]). In principle, this would require a proof obligation that guarantees that the axiom can be dropped, but in this case the proof is trivial, since \([ax2']\) coincides with the instance of \([\textit{genC1}]\) relative to operation \textit{test}.

Let us consider now axiom \([ax5]\) of class GPDB. In this case one has to take into account that the TRIO item \textit{cyclic\_acq} has become the operation \textit{get\_measure} and that when the latter ends the information sent back is described in a more detailed way, since a data structure made up of three fields (\textit{status}, \textit{oper\_mode} and \textit{acc\_perm}) is used. Therefore, axiom \([ax5]\) is rewritten as follows:

\[
(\ \text{get\_measure}(i).\text{end\_success} \quad [ax5'])
\wedge\ \text{Past}(\ \text{get\_measure}(i).\text{inv\_received} \land \text{get\_measure}(i).\text{device} = \text{dev}, \text{Tmax})
\wedge\ \text{MC\_address}(\text{dev}, \text{MC\_ad})
\wedge\ \text{WithinP}(\text{status}(\text{MC\_ad}, \text{dev\_s}, \text{om}, \text{a\_p}), \text{Tmax}))
\rightarrow
(\ \text{get\_measure}(i).\text{brief\_status}.\text{status} = \text{dev\_s}
\wedge\ \text{get\_measure}(i).\text{brief\_status}.\text{oper\_mode} = \text{om}
\wedge\ \text{get\_measure}(i).\text{brief\_status}.\text{acc\_perm} = \text{a\_p})
\]

In other cases, the TC description may contain axioms that do not exist in the specification. Such axioms typically describe some lower-level behaviors not previously taken into account.

For example, operation \textit{variation} has an input parameter, named \textit{calibrations}, composed of five fields (\textit{calibID}, \textit{date}, \textit{zero\_error}, \textit{span\_error} and \textit{lin\_eq}) used to send some calibration data to the IMS. A new axiom is introduced to specify that when calibration data are sent all the information must be defined.
\[
\text{variation}(i).\text{calibrations}(l).\text{calibID} = \text{cal} \quad \quad \quad \quad \quad \text{[axN]}
\]
\[
\rightarrow \exists \ d, \ z_e, \ s_e, \ \text{lin_eq} \\
\quad \quad ( \ \text{variation}(j).\text{calibrations}(l).\text{date} = \text{d} \\
\quad \quad \quad \land \ \text{variation}(j).\text{calibrations}(l).\text{zero_error} = z_e \\
\quad \quad \quad \land \ \text{variation}(j).\text{calibrations}(l).\text{span_error} = s_e \\
\quad \quad \quad \land \ \text{variation}(j).\text{calibrations}(l).\text{lin_eq} = \text{lin_eq})
\]

This level of detail was not taken into consideration in the specification, but is suitable for an architectural description.

As a last example let us consider the choice, discussed during step 1, of using a field-bus to implement the communication between GPDB (currently represented by CORBA Entity classes DataRep and Gateway) and the field devices. Moreover, let us suppose that one wants to state that values coming from the devices (i.e., whenever \(\text{ctrl}\) is true) represent

1. the results of a test/command issued by IMS via Gateway, which must be sent within \(T_1\) time units to IMS, or
2. the results of a cyclic data acquisition performed by IMS via DataRep, which must be sent within \(T_2\) time units to IMS, or
3. the variations occurred in some device that must be notified to IMS within \(T_3\) time units.

The above property involves several different components of the architectural description of the system and thus is formalized by the following axiom in the \textit{Environment} class Maintenance System, which represents the whole application:

\[
\text{MeasuringChannels}[j].\text{ctrl} \\
\rightarrow \exists \ i, \ d \\
\quad \quad ( \ ( \ \text{WithinF}(\text{Gateway.test}(i).\text{end_success} \land \text{Gateway.test}(i).\text{device}=\text{d}, \ T_1) \\
\quad \quad \quad \lor \ \text{WithinF}(\text{Gateway.command}(i).\text{end_success} \land \text{Gateway.command}(i).\text{device}=\text{d}, \ T_1) \\
\quad \quad \quad \lor \ \text{WithinF}(\text{DataRep.get_measure}(i).\text{end_success} \land \text{DataRep.get_measure}(i).\text{device}=\text{d}, \ T_2) \\
\quad \quad \quad \lor \ \text{WithinF}(\text{IMS.variation}(i).\text{inv_received} \land \text{IMS.variation}(i).\text{device}=\text{d}, \ T_3))
\]
\[
\land \ \text{GPDB.MC_address}(\text{d}, \ j)
\]

where \(MC\_address\) is a predicate binding each instance of a device (index \(j\)) with its symbolic name, used by IMS (variable \(\text{dev}\)).

Furthermore, other axioms, not reported here, ensure that each time \(\text{ctrl}\) is true only one of the above operations occurs.

Some of the transformations of this section can be done automatically; however, the intervention of the designer is often crucial. For example, once we have recognized that \textit{test}\_\textit{request} corresponds to invocation of operation \textit{test}, transformation of axiom [ax1] into [ax1'] is straightforward, and could be done automatically. The same holds for elimination of axiom [ax2]. Modification of axiom [ax5], instead, relies heavily on the intuition that the designer has of the low-level details of method invocation, and would be difficult to be done automatically. Conversely, the last two axioms do not come from the specification and have been added by designers thanks to their knowledge of the system.

5 BENEFITS AND LESSONS LEARNED

During the design of the TC language and methodology, and the development of the MS application, a number of issues came up providing a qualitative assessment of the approach taken, the achieved benefits, the lessons learned, and remaining problems and challenges. This section discusses the most significant aspects of our assessment, based both on the experience gained during the OpenDREAMS-II project and on the subsequent research efforts.
The development of the MS application confirmed once more that rigor and formality are fundamental, at least when dealing with the most critical aspects of an application, in order to understand and manage complex systems' specification and architecture design. Safety-critical (distributed) systems are natural candidates for exploiting formal methods techniques because of their inherent complexity. During the process of formally specifying some aspects of the MS application new problems, ignored by the first informal description, aroused (e.g., formalization of services pointed out flaws in their definitions [OpenDREAMS-II 1998b]).

One of the greatest benefits of formal methods lies in the ability to formally verify correctness of the design against the specification before the system is actually implemented. In our approach, formally proving that a TC architecture is consistent with the associated TRIO specification can be carried out using the PVS system [Owre et al. 1992]. The prototype of a correctness prover -or disprover- based on the translation of the TRIO formalism into PVS [Alborghetti et al. 1997, Gargantini and Morzenti 2001] has been successfully used to demonstrate the correctness of the architecture of a test application with respect to its specification. Moreover, using TC to formally define the application architecture allowed us to apply the TRIO techniques and tools [Mandrioli et al. 1995, San Pietro et al. 2000] to verify that the design does not contain inconsistencies or to derive real-time functional test cases from specifications. These issues are not discussed in this paper but the interested reader can refer to [Rossi 2002] in which TC verification issues are discussed.

A major benefit coming from coupling formal methods with an open, standard middleware architecture such as CORBA has been the increase of the level of abstraction at which designers have to work. For example, many communication issues among the different components of the application, which had to be explicitly taken into account in our previous experiences, turned out to be formalized by the TC language and thus they could be used in an off-the-shelf fashion.

Another benefit related to our approach consists of its ability to follow in a smooth way the technological changes that may occur. In fact, since the beginning of our research, the OMG’s specifications have evolved: new services were introduced and existing specifications have been modified. Our approach allowed us to follow the specification evolution in a fairly natural way: whenever new features were added by the OMG, TC simply required the definition of new (built-in) axioms and possibly the use of stereotyping mechanisms, without having to modify the core of the language. For instance, we originally started with the standard non-real-time CORBA and thus strict timing requirements were not taken into account. Once the OMG’s issued the Real-Time CORBA specification [OMG 2002b] we were able to use the features presented therein in a quite direct and straightforward way, with a minor adaptation of our methodology (see for instance Section 4.3.1).

Moreover, even if this paper focused on CORBA-based architectures, the same approach can be adapted and applied to other object-oriented middleware such as DCOM [Eddon and Eddon 1998] or Java/RMI [Pitt and McNiff 2001]. In fact, by defining our own syntax for Interface classes rather than modifying OMG’s IDL, we could add features to our interface language and translate it into a variety of other Interface Definition Languages. Finally, TC and its related methodology (possibly with semantic extensions, introduced by means of new stereotypes) can be applied to a variety of other middlewares, such as event-based [Cugola et al. 2001] and peer-to-peer ones [Oram 2001].

So far, the work done has been carried out by a strongly integrated and long term cooperating group (composed of people from the academia and the industry), who share a common background in the area of formal methods. Thus, it remains to be verified if similar benefits can be achieved by industrial designers who have not been so deeply involved in the
development of the methodology. This is a fairly typical and still daunting problem for the diffusion of formal methods in the practice of industrial projects [Ciapessoni et al. 1999].

Another, somewhat related, problem that could hamper a wider diffusion of our methodology is represented by the currently available tools supporting our approach both when specifying/designing an application and when performing verification activities. Currently, the prototype of a graphical interactive editor supporting the documentation of all phases, from requirement specification to architectural design, is available, while no verification tool specifically tailored for TC is currently available. Thus, the work done has been carried out by translating TC into TRIO and then using the available (semantic) tools. This approach required a high level of expertise from the users and may not be viable without a deep understanding of TRIO and its semantics.

Finally, the methodology has been developed hand-in-hand with the MS application and it can be seen as coming from it. Beside the MS application, TC and its related methodology has been applied to other smaller cases (e.g. [Morzenti et al. 1999]). It is likely that by applying our methodology to more case studies, one could get additional useful insights to revise and refine some of its finer points.

6 RELATED WORKS

The TC language and methodology presented in this paper combine in a novel way concepts and experiences derived and adapted from a number of different (and, until now, mostly unrelated) research fields such as:

- Architecture Description Languages;
- extensible formalisms and formal methods;
- object-oriented design methodologies;
- middleware and architectures for distributed applications.

Therefore, our approach shares some background ideas with many of the research efforts in the above fields, but it distinguishes from them essentially in that it combines and integrates those ideas and applies them to the domain of (critical) CORBA-based, distributed applications.

The rest of this section analyzes the relationships of our approach with the results obtained in the areas mentioned above. In addition, it puts the TC language/methodology in the context of the broader research effort carried on at Politecnico di Milano of which this work is part.

Historically, the combination of formal methods with a stepwise refinement process that allows one to proceed smoothly from high level specifications down to architectural design and possibly even to final implementation dates back to early algebraic and logic approaches such as Larch [Guttag and Horning 1993] and B [Abrial 1996]. However, the application of such approaches has been limited to traditional, general purpose (in most cases sequential) systems, while distributed, real-time systems were not taken into account.

Architecture Description Languages (ADLs), such as Darwin [Magee et al. 1995] and Wright [Allen and Garlan 1997], allow one to describe software systems in term of communicating components and, in the case of Wright, connectors. However, most of them focus on the static structure of the system rather than on the components’ dynamic behavior, which is often dealt with separately (possibly with a different formalism) [Kramer and Magee 1997]. In addition, the formalisms chosen to model components' behavioral aspects are mostly CSP-like [Hoare 1985], which are unsuitable for describing time-related properties. TC, instead, allows the designer to formally describe both the structure of a system and the behavior of its
components. Being TC based on a temporal logic such as TRIO, time-related properties can be described in a very natural way.

The Rapide ADL [Luckham et al. 1995] takes into account both structural and behavioral aspects and allows one to express temporal relationships between events in the system [Luckham 1998]. With respect to Rapide, the TC is more flexible in time modelling, as it can deal with both dense and discrete time domains.

In addition, [Di Nitto and Rosemblum 1999] showed that most ADLs are not flexible enough to allow designers to take into account the specificities of the underlying (middleware) technology such as CORBA. The main reason being that they tend to be limited in scope, and they offer only a small number of constructs to avoid complexity explosion in the simulation of the architecture. TC, instead, provides CORBA specific features that allow designers to explicitly take into account the main aspects of the target technology. Moreover, TC offers extension mechanisms that allow designers to formally introduce new concepts by means of the existing ones. Thus, TC can be used to describe the system at different levels of abstraction. For example, CORBA specific features are modeled in TC by defining the basic concepts (e.g., operations, CORBA objects, etc) in term of TRIO, while higher-level concepts (services, frameworks, etc.) are defined in terms of the basic ones.

Another distinguishing feature of our approach with respect to others such as Darwin [Magee et al. 1995], Durra [Barbacci et al. 1993] is being tailored towards SCSs, which are mostly demanding in terms of reliability—and often are hard real-time systems. Such an orientation, however, does not affect the whole method, which in large part is well suited for general distributed applications based on CORBA; only the final step, which exploits typical services and frameworks, is specialized towards this application domain. In fact, we also applied the method to other, non-SCS applications [Morzenti et al. 1999].

Stereotyping is used in many approaches to introduce specific constructs in a language. For example, [Robbins et al. 1998] uses stereotyping to extend UML with ADL-specific concepts, without providing a formal definition of such concepts; [Kaveh and Emmerich 2001] introduces stereotypes to formally model synchronization primitives and thread management in CORBA architectures to detect deadlock through model checking. However, the approach of [Kaveh and Emmerich 2001] does not consider timing requirements. In our approach, instead, stereotypes are formally defined and they can also express timing properties.

Another peculiar feature of our approach is that TC couples the language for describing high-level system architectures with a development methodology, which starts from system requirements. Therefore, TC goes beyond ADLs, which lack a methodology to determine the system architecture from the requirements.

With respect to other general purpose, object-oriented (OO) methodologies, the TC approach is explicitly targeted to CORBA, has formal foundations, and deals rigorously and quantitatively with time. Each of these features can be found singularly in literature but no one combines all of them in a single development method. Methodologies and notations such as [Booch 1994, Booch et al. 1996, Rumbaugh et al. 1991, Jacobson et al. 1999, Maher et al. 1996] do not specifically address the issues of Object-Oriented Analysis/Development over CORBA, nor (with the exception of the OCTOPUS notation) those of real-time systems. Moreover, they do not allow a formal description of requirements since they lack a rigorous underlying mathematical model, even though some work has been carried out to couple these methodologies with formal specification languages [Lano 1996, Lavazza et al. 2001]. ROOM (Real-time Object Oriented Modeling, [Selic et al. 1994]) and UML for real-time [Douglass 2000] (which owes many of its concepts to ROOM) are notations tailored towards real-time
systems; however, these notations are mostly informal and they do not address the specificities of middleware-based distributed systems.

To fill this gap, the OMG, which now supports UML [OMG 2003], has added a CORBA profile to UML [OMG 2002c]. It also began to support model-driven application development [OMG 2001c], and has defined a metamodel to describe software development processes [OMG 2001d]. However, OMG's efforts in the field of CORBA-targeted development processes are still largely informal, and they do not tackle real-time issues. Notice also that such efforts began significantly after, not only the beginning, but even the first results of our research along this approach [Coen-Porisini and Mandrioli 2000, Coen-Porisini, et al. 2000].

Finally, [Neema et al. 2002] define a modeling language, called Adaptive Quality Modeling Language (AQML), that is used to describe the requirements concerning the management of policies within a Quality of Service (QoS) adaptation layer built on top of middleware such as CORBA. The AQML approach is oriented towards modeling and simulating the QoS adaptation and the underlying middleware infrastructure. The TC approach is more general because it is geared towards modeling all aspects of a CORBA-based system, from middleware infrastructure to application logic. In addition, AQML cannot express time-related properties, which is a distinguishing feature of TC.

The TC language and methodology presented in this paper are part of a long term research going on at Politecnico di Milano, whose ultimate goal is the development of sound, proven correct, real-time, distributed applications based on the CORBA platform. [Coen-Porisini and Mandrioli 2000] reports on the first results of the research, which were based on the initial, non real-time, CORBA standard and on early case studies; such first results were still rather informal and the TC language was not yet defined. [Coen-Porisini et al. 2000] provides a first description of both the TC language and methodology but due to typical conference proceedings limitations, it does not discuss many aspects of the methodology and lacks a comprehensive presentation of the TC language.

Thus, this paper extends and revises [Coen-Porisini et al. 2000] by including many more technical details and a more thorough description of both the methodology and the case study. Moreover, this paper also reports on further results that have been achieved more recently, while remaining focused on the key aspects of formally modeling application requirements and its architecture and on the methodology to derive the latter from the formers.

Other related results are documented in separate reports and papers: [Rossi 2002] discusses the formalization of real-time aspects, which has proceeded in parallel with the evolution of the standard and the development and availability of real-time ORBs, and [Pradovera 2001] reports on an experimental validation of the real-time features of TAO [Schmidt et al. 1998]. [Rossi 2002] deals also with the verification aspects of the TC approach, with strong attention to hard real-time constraints. It introduces a two-layered model suitable for describing and verifying real-time CORBA-based applications with different levels of detail: the lower-level (more detailed) layer takes into account the ORB and the operating system elements such as RTPOAs and schedulers. The low-level mechanisms of the Fault-Tolerant CORBA specification [OMG 2002a], instead, have yet to be included in the TC model.

Since revision 2.5, Fault-Tolerant CORBA is part of the CORBA/IIOP specification [OMG 2001a]. The FT CORBA specification does not affect the TC methodology since, from the methodological point of view, what is relevant is that an object is reliable, not how this is achieved. A formalization of the FT CORBA specification, however, would lay the basis to prove that the mechanisms described therein are suitable to obtain the desired degree of reliability and, in particular, that FT CORBA enforces the semantics of the «reliable»
stereotype. This issue, however, relates to the problem of modeling and verifying the CORBA platform, which is outside the scope of this paper.

7 CONCLUSIONS

This paper proposed and illustrated a formal method to develop distributed applications based on CORBA. The method exploits the object-oriented logic language TRIO and drives the designer to obtain a complete CORBA architectural design through a smooth sequence of steps starting from the specification of the application requirements. Thus, the method enjoys the typical benefits of formality, i.e., rigor and precision, both in specification and in verification, and the possibility of using powerful tools (e.g., to generate (semi) automatically test cases for the implementation). In particular, since the semantics of both application specification and architectural design is expressed in terms of logic formulas one can, at least in principle, prove the correctness of the design as a typical logical implication.

This paper focused on the essentials of the method. The reader is referred to the bibliography for a more thorough and detailed exposition. In particular, [Pradella 2000] describes the method and the application case study in full detail. The fundamental issue of managing real-time aspects in CORBA-based systems, not considered in this paper, is the objective of [Marotta et al. 2001] and [Rossi 2002] where the real-time extension of CORBA [OMG 2002b] is analyzed and formalized and it is shown how to build -potentially- guaranteed real-time applications on top of it. The issue of verifying the correctness of the architectural design against application requirements is addressed by [Rossi 2002].

In our approach, we chose not to modify in any way the definition of CORBA (e.g., we do not propose any formal extensions to IDL). Instead, we decided to preserve its basic features, coupling them with a formal definition. This approach should not be seen as an alternative to existing non-formal, non CORBA-oriented methods such as UML; rather, it is well suited to augment, and be integrated with, several existing informal practices [Ciapessoni et al. 1999].

The evolution of CORBA (and related research) over the time has also confirmed the validity of our approach, which stresses the importance of reliability and time predictability in a CORBA architecture. The efforts of OMG towards both supporting more precise modeling formalisms (UML is a first step in this direction, although very incomplete) and adding features specifically meant to improve reliability and time predictability of the platform (e.g., Fault-Tolerant and Real-Time CORBA) proved the soundness of the goals initially set by the OpenDREAMS projects, which this research was part of.

Finally, we expect to produce a suite of TRIO/TC tools that will be easily accessible and widely usable in an industrial environment. In fact, TRIO is currently supported by a suite of (prototype) tools that cover many aspects of formal verification. Thanks to the distinguishing feature of TRIO that complex operators are built on top of a basic, simple building block (operator Dist) these tools have smoothly evolved over the time to include high-level concepts (e.g., states and events) that were not present in the original version of TRIO. In a similar way, new evolutions are planned to include TC elements. In fact, since the semantics of TC is entirely given in TRIO, TC tools can be built upon TRIO ones by translating TC architectures into TRIO specifications (i.e., adding the necessary built-in axioms every time that a TC-specific construct is used). However, this approach might not be very efficient, and therefore future work will aim at optimizing TRIO tools to better deal with TC-specific concepts.

ACKNOWLEDGMENTS

As already mentioned, this work has been done during the ESPRIT projects OpenDREAMS-I and II. We acknowledge the cooperation of all the members of the project: Alcatel CRC, ENEL, EPFL, ISR, Poet, and Teamlog.
We also thank the reviewers for their comments, which helped us improving this paper.

REFERENCES


The reader interested in acquiring the OpenDREAMS reports should contact the authors.

18


IEC - IS - 1158-2. Field Bus standard for use in industrial control system Physical layer specification and service definition.


OPENDREAMS CONSORTIUM, Final Report, Deliv. WP0/T0.1-ALCT-REP/R01e


