Object-Oriented Concurrent Reflective Architectures *

Satoshi Matsuoka Takuo Watanabe Yuuji Ichisugi
Akinori Yonezawa
Department of Information Science, The University of Tokyo†

Abstract

Reflection provides the abilities to reason and alter the dynamic behavior of computation from within the language framework. It is a practical scheme that offers a new perspective in constructing a malleable, large-scale system such as programming languages, operating systems, and window systems. Reflection is more beneficial in object-oriented concurrent/distributed computing, where the complexity of the system is much greater compared to sequential computing; this has been demonstrated through recent works in *Object-Oriented Concurrent Reflective (OOCR)* architectures. From a structural perspective, we can categorize OOCR architectures into *individual-based*, *group-wide*, and their hybrid (*hybrid group*). Through the introductive categorization of various OOCR architectures based on our previous and current work, we explore the characteristics as well as benefits and limitations of each.

1 Introduction to Reflection and Reflective Systems

Reflection is the process of reasoning about and acting upon the system itself[23, 15, 34]. In a conventional computing system, the *subject matter* of computation is external to the system, i.e., the computation is performed on data that represent or model entities that are external to the system. A *reflective system* takes a step further by allowing the subject matter be the system itself, i.e., computation can be performed on the system itself in a uniform way[15]. A reflective system thus 'opens up' the system by providing an appropriate abstraction on the internal (often implementational) detail of itself to the user program. Contrary to the misconception that 'reflection' is some difficult-to-understand, not-too-useful philosophical jargon, it is a practical scheme that offers a new perspective in constructing a malleable, large-scale system such as programming languages[14], operating systems[32], and window systems[22]. One can make an analogy of reflection to

*Presented at the Workshop on Object-Based Concurrent Programming, Geneve, Switzerland, July, 1991. (Proceedings to be published as Mario Tokoro et.al. eds., a volume in Lecture Notes in Computer Science, Springer-Verlag, 1992)

†Physical mail address: 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan. Phone 03-3812-2111 (overseas +81-3-3812-2111) ex. 4108. E-mail: {matsu,takuo,ichisugi,yonezawa}@is.s.u-tokyo.ac.jp

inheritance, which was derived from AI and incarnated into object-oriented programming languages, and now serves important roles in methodologies for system construction of user interfaces, operating systems, and database systems. We believe that, in the same manner, reflection can offer another clear architectural perspective, especially for the dynamic aspects of the system where inheritance is known to be somewhat ineffective.

To expose, or *reify* its internals, a reflective system contains data that represents or models the metalevel structural and computational aspects of itself within itself. Such data is said to be *reified*, and must be dynamically self-accessible and self-modifiable by the user program. Furthermore, when the user modifies such data, the modification must be 'reflected' to the actual computational state of the user program. This requirement is termed as *causal-connection*, and is satisfied by all reflective systems. When the requirement is satisfied, the meta-level data that provides the abstraction is called the *Causally-Connected Self-Representation(s) (CCSR)*. For practical purposes, the meta-relation is often that of implementation. In such systems, causal connection is automatically guaranteed. Although modifying the implementation may seem like a dangerous endeavor, it has been shown that modification can be confined to the local portion of the user program by some appropriate *Metaobject Protocol*, a scheme pioneered in CLOS[14].

Another requirement for reflective systems is that self-access and self-modification must be possible in a way that is little different from the base-level programming. In other words, appropriate abstraction of the internal program state must be provided by the system so that the user can perform the meta-programming of the system in a clean, concise manner that does not deviate significantly from 'normal' programming. For example, in 3-Lisp[23], one can program reflective computation merely by defining *reflective procedures* whose means of definition differ little from that of normal procedures.

2 Reflection in Object-Oriented Concurrent Systems — The Benefits

The benefits of *computational reflection* in sequential languages have been identified by many researchers: Previously, programming language functionalities such as debugging facilities, interfaces to external world, and exception handling were added to languages in an ad hoc and inflexible manner. In contrast, languages based on reflective execution models are powerful in providing linguistic mechanisms to use the above functionalities in a uniform and flexible manner via clean access of the abstraction on how the system is implemented[15, 14].

We claim that reflection is more beneficial in concurrent computing, for a concurrent system embodies multitudes of aspects that do not arise in sequential computing: scheduling, communication, and load-balancing, among numerous others. Those aspects are still considered to be mostly outside the scope of programming languages, and their control is only available in a fixed, ad-hoc fashion, with little possibility for user extensibility. This is a not favorable, since concurrent architectures are much more complex as compared to sequential ones, and the system must be *open* to structured dynamic extensions and/or modifications for adapting to new problems and environments.

Here, as were pointed out in [34] and [24], constructing a concurrent system with a reflective architecture could be beneficial to encompass such aspects within the programming language framework for the reasons mentioned above. The dynamic optimization is especially effective in a concurrent system because (1) as mentioned, the reflective architecture allows the programmer to control the computational aspects of concurrent system dynamically[1], and (2) both the reflective computation and the base-level computation can be performed in parallel[34, 17, 29].

3 Object-Oriented Concurrent Reflective (OOCR) Architectures

With the aforementioned observations and principles, our research proposed the use of reflection in the framework of *Object-Oriented Concurrent Programming (OOCP)*, with various *Object-Oriented Concurrent Reflective (OOCR)* architectures: languages ABCL/R[29] and X0/R[18] were based on the *individual-based architecture*, ACT/R[30] was based on the *group-wide architecture*[2], and the evolution of ABCL/R, the language ABCL/R2[17], was based on the *hybrid group architecture*. Each language differed in what meta-lingual features can be realized through reflection, i.e., what features particular to concurrent OO computing, such as object monitoring and migration, could be cleanly encompassed within the programming language framework. There, the evolution of the languages can be understood for the difference in the characteristic of their reflective architectures.

Although there have been works that discusses various aspects of reflection and classifies metalevel/reflective architectures from different viewpoints[16, 28, 6, 24], so far to our knowledge no work has existed that generally discusses reflection in the context of concurrent/distributed object-oriented languages. Nor, has there been work that classifies architectures dependent on characteristics that are *particular to object-oriented concurrency*, i.e., not becoming manifest in sequential object-oriented languages. It is our premise that the need for such a discussion and classification is becoming omnipresent — for as we argued above, the power of reflection can be best utilized in concurrent computation. Furthermore, we claim that in order for reflection to be put to practical use in concurrent/distributed systems, we need to analyze and evaluate the merits and weaknesses of each, and synthesize a new architecture suitable for practical use.

First, the two aspects of CCSR in OO languages are the (1) *structural aspect*, indicating how objects or group of objects in the base-level and the meta-level are constructed and related, and the (2) *computational aspect*, indicating how meta-level objects represent the computation of the base-level objects[6]. Altogether, appropriate construction of the *meta-level architecture* dictates what structural and computational aspects can be reflected and operated upon — this we believe is more prevalent in concurrent systems, for factors such as scheduling must be given appropriate meta-level abstractions in order for them to be effectively and efficiently accessed from the base-level.

The OOCR architectures can be categorized from the structural aspect, chiefly dependent on whether the architecture has the notion of *metaobjects*, which Pattie Maes introduced for sequential object-oriented language 3-KRS in [15]. It is often argued that providing explicit metaobjects in the architecture is conceptually natural with respect to

[1] One of the interesting applications we will note is to change the scheduling policy of object execution dynamically [3] in distributed simulation based on the Time Warp Scheme[13].

[2] The original idea of the group-wide architecture was mentioned in the last sections of [29] and [34]

the uniformity principle, "everything is an object." But, what do we really mean by a 'metaobject'? It is pointed out that there has been great confusion on the usage of the term[7] — we first attempt to provide conceptual distinction in the terminologies.

A *meta-level object* is any object that resides in the meta-level of the object-level computation. *Metaobjects* are objects that reflects the structural and computational aspect of a *particular* base-level object; thus, a metaobject is a meta-level object, while the converse is not necessarily so. The relationship between the metaobjects and the base-level object they represent could either be one-to-one or many-to-one, i.e., there could be multiple metaobjects collectively 'representing' a single base-level object[10, 21]. The metaobjects must be unique to a single object in the base-level, that is, they are not shared.

There are several important criteria that metaobjects must satisfy that distinguish it from other meta-level objects:

1. *Identifiability of Metaobjects* — a base-level object must be able to identify its metaobjects intrinsically.

2. *Uniqueness of Metaobject Denotation* — metaobjects must be able to intrinsically and uniquely identify the base-level object it denotes.

3. *Causal Connection* — a base-level object and its metaobjects must be *causally connected.*

4. *Governing of Computation* — the metaobject must dynamically govern the computation of the object it denotes or represents.

By these criteria, the 3-KRS and ABCL/R 'metaobjects' are metaobjects as we describe later on, whereas the CLOS 'metaobjects' are meta-level objects in our terminology. Class objects in class-based languages such as ObjVLisp[4] are not metaobjects either, for (1) the denotation is not direct and/or not unique, and (2) classes are not "meta" in a computational sense[6], that is, they do not actually govern the dynamic aspect of object computation.

Given the distinction of metaobjects and meta-level objects, a reflective object-oriented system can basically be constructed in two ways:

- **Individual-Based Architecture:** Each object in the system has its own metaobject(s) which govern(s) its computation. By 'individual-based' we mean that an individual object is the unit of base-level computation that has a meaningful CCSR at the meta-level.

- **Group-wide Architecture:** In this architecture, a group of meta-level objects comprises the meta-level. The behavior of an object is not governed by metaobjects; rather, the collective behavior of a *group* of objects is represented as the coordinated actions of a group of meta-level objects, which comprise the *meta-group*. By 'group-wide' we mean that the entire object group is the unit of base-level computation that has a meaningful CCSR at the meta-level (as a meta-group); thus, there are *NO* intrinsic meta-relationships between a base-level object and a particular metaobject at the meta-level.