

Co-evolution of Object-Oriented Software Design and Implementation

Theo D'Hondt – Kris Devolder – Kim Mens – Roel Wuyts
Programming Technology Lab – Vrije Universiteit Brussel

Abstract: Modern-day software development shows a number of feedback loops between various phases in its life cycle; object-oriented software is particularly prone to this. Whereas descending through the different levels of abstraction is relatively straightforward and well supported by methods and tools, the synthesis of design information from an evolving implementation is far from obvious. This is why in many instances, analysis and design is used to initiate software development while evolution is directly applied to the implementation. Keeping design information synchronized is often reduced to a token activity, the first to be sacrificed in the face of time constraints. In this light, architectural styles are particularly difficult to enforce, since they can, by their very nature, be seen to crosscut an implementation. This contribution reports on a number of experiments to use logic meta-programming (LMP) to augment an implementation with enforceable design concerns, including architectural concerns. LMP is an instance of hybrid language symbiosis, merging a declarative (logic) meta-level language with a standard object-oriented base language. This approach can be used to codify design information as constraints or even as a process for code generation. LMP is an emerging technique, not yet quite out of the lab. However, it has already been shown to be very expressive: it incorporates mechanisms such as pre/post conditions and aspect-oriented programming. We found the promise held by LMP extremely attractive, hence this paper.

1. INTRODUCTION

Recent times have seen a consolidation of methods and tools for software development in production environments. A good number of de facto standard tools have emerged, not the least of which is the *Unified Modelling Language* in some commercial incarnation. At last, one could say, we have come to know the process by which the design and implementation of a complex piece of software is charted. Objects have been widely accepted, the programming language babel seems more and more controlled by the emergence of Java and previously untamed regions of information technology, such as distribution, coordination and persistence, are starting to become daily fare in software applications.

So why is software development still arguably the least predictable of industrial processes? Why can comparable software projects, executed by development teams with comparable skills, not be planned with comparable margins of error? Why is our appreciation of the software development process still flawed, even after the introduction of all these new techniques and tools?

For some time now, grounding the development of software in a programming language has proved not to be scalable. This led to the notion of *software architectures* as a collection of techniques to buttress this development process, particularly in those places where programming languages or tools fail to capture the macroscopic structure of the system that needs to be built. Close to the programming language technology itself, we find the well-understood *framework* approach; at a more abstract level we find techniques built on various kinds of *patterns* and *contracts*.

The latest landslide in this fight for control over software complexity is the emergence of *component* technology. At a time when commercial component toolkits such as *Enterprise Java Beans* are proposed as the solution to our problems, we do well in realizing that the advent of components amounts to an acknowledgment of defeat. In fact, by accepting this technique of decomposition into static components, we have come full circle and reinvented data abstraction. The task of making components cooperate is not any better understood as any of the numerous software building strategies we have taken under consideration these past 20 years.

An important step in understanding this partial failure is the insight that software is fluid. It is in constant evolution under the influence of ever changing conditions; software development is sandwiched between a technology that is evolving at breakneck speed, and requirements that must follow the economic vagaries of modern society. In this, the commercial product called *software* is unique; the closest professional activity to that of software developer is that of composer in 18th century Europe. At that time, relatively widespread knowledge of harmony or counterpoint made up for the necessary skills to use and reuse fragments of sophisticated musical artefacts. For instance, [LD96] offers insight on how *invention*, a term borrowed from rhetoric, drives composition according to a process which bears a striking resemblance to building complex computer applications. Unfortunately, equivalent skills needed to master a software artefact are today in far more limited supply than 250 years ago.

This contribution is a synthesis of recent work performed by various people within our lab in addressing this need for more control over the evolution of software. Although in the past, a significant amount of work focussed on the need to *document* evolution [SLMD96, CL97], to build *conflict detection tools* [KDH98] and to formalize the evolution process [TM9], we will concentrate here on an emerging approach for *steering* evolution. This is very recent work and as such has only resulted in experiments and prototypes. We feel however that it is sufficiently mature and promising to be presented here as a whole. In the bibliography we limit ourselves to a number of key documents¹ describing these activities; these in turn contain a much more comprehensive list of references.

We have chosen to use the term *co-evolution*, implying that managing evolution requires the synchronisation between different layers (or views) in the software development process. We will therefore dedicate the next section to an analysis of this statement. Next, we propose a concept called *Logic Meta Programming* (or LMP for short) as a development framework in which to express and enforce this synchronisation process. Another section of this paper will be used to introduce LMP and to situate it in the broader context of software development support. Finally, several experiments with LMP will be presented in evidence of its applicability. We will discuss using LMP as a medium for supporting *aspect oriented programming*, for enforcing *architectural concerns* in an object oriented programming environment and to express constraints on the protocol between a collection of interacting

¹ Available via <http://prog.vub.ac.be>

software components. This is by no means a complete coverage of LMP, nor even of the experiments conducted at our lab; we feel however, that it provides sufficient insight in the applicability of LMP to the co-evolution of software, while avoiding exposing the reader to too much detail.

2. SYNCHRONIZING DESIGN AND IMPLEMENTATION

Currently accepted procedures in the development of software involve the consideration of several views. In descending order of abstraction one encounters *requirements capture, analysis, design, implementation, documentation* and *maintenance*. Not quite by accident, this also happens to constitute an ordering according to increasing level of detail, albeit not a continuous one. There is in fact a kind of watershed between design and implementation, which commits the developer to a level of detail that is very hard to reverse.

Consider an example describing a simple management hierarchy. At best this is captured at the class diagram level by an arity constraint, although the more subtle aspects such as the required absence of cycles in the hierarchy graph can only be expressed by an informal annotation:

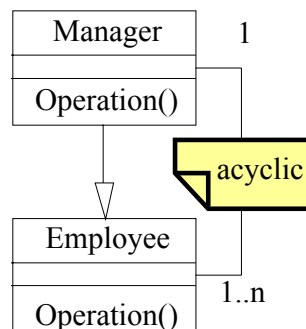


Figure 1: a manager/employee hierarchy

It can be seen that as we transit from the original requirements to the design, we replace abstract concepts by more concrete ones. The same holds for the implementation: in our simple example the arity constraint might be replaced by a precondition in a mutator function while the acyclicity constraint, if implemented at all, gives rise to some consistency maintenance code. On the other hand, we see that this decrease of abstraction is compensated by an increase in the level and amount of detail.

This observation holds in general and will be viewed as trivial by most software developers. However, we do well in analysing this transition from abstraction to detail as we descend through the various levels in the lifecycle of a software application. Typically, the amount of energy that needs to be applied increases with the level of detail; so does the need for technical skills. This generally makes an implementation artefact more valuable than a design artefact. Also, any ultimate defect in respecting the original requirements is detected at the lowest level, i.e. the implementation.

Initially, the development of a software application is achieved by this progression through the various abstractions: requirements, analysis, design and implementation. However, once the implementation has reached the production stage, the tangible aspects of the prior stages are at best used as documentation in order to boost understanding of the actual code; at worst they become obsolete. This is a well-known phenomenon: under the pressure to bring software to market in the face of competition, or to correct flaws under the threat of contractual penalties, the management of evolving software all too often degenerates into updating implementations. Various directions have been explored to improve this situation: in general they imply some re- engineering activity applied to evolving implementations in order to extract abstractions and update e.g. design documentation. Hardly anyone uses an approach where design concerns drive the implementation process; programming environments that explicitly constrain the developer to design decisions are hard to find. Popular languages like Java evolve, but they evolve towards a more sophisticated type system: boosting genericity is a technical issue and hardly qualifies as support for e.g. architectural concerns.

It is our conjecture that during the development process, the concretisation of abstract concerns should not consist of some kind of erosion. On the contrary, any relevant feature should be kept available in any of the ulterior phases. We will in particular concentrate on the synchronisation between *design* and *implementation*. For the sake of this discussion we will discard requirements that cannot be expressed as explicit design directives. Our ambition is to augment an implementation such that it becomes a strict superset of its design; design can be extracted from an implementation by ignoring details; design can be interpreted by the programming environment and therefore enforced. We propose an approach called *Logic Meta Programming*, (or LMP for short) which will be described in the now following section.

Consider as an example a change in the manager/employee example where a decision to introduce *workforce pooling* results in the arity constraint to be changed into:

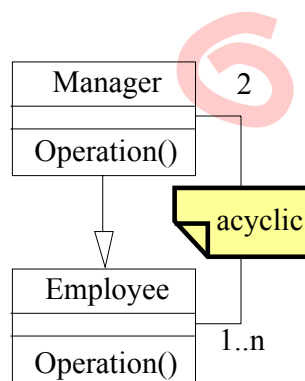


Figure 2: a manager/employee hierarchy with workforce pooling

The program code will probably need to be significantly changed with hardly an explicit link to the original arity constraint. We would prefer it to be explicitly present in the implementation as some kind of enforceable declaration, formatted in the proposed LMP-paradigm.

3. LOGIC META PROGRAMMING

Logic Meta Programming, or LMP for short, is the name we have given to a particular flavour of multi-paradigm programming. The starting point for LMP is an existing programming environment that is particularly suited for engineering large software systems. In this contribution, we have limited ourselves to a Java-based environment and to a Smalltalk-based environment. Next, we augment this environment by a declarative meta layer of a very particular nature. In the case of Java, i.e. a statically typed language, this meta layer might be implemented as a pre-processor or even an extension of the Java compiler itself. In the case of Smalltalk, it requires the installation of a number of class methods and class variables in the standard Smalltalk hierarchy. We are interested in a declarative approach; it seems intuitively clear that design information, and in particular architectural concerns, are best expressed as constraints or rules. The acyclicity constraint from the manager-employee example on the previous page seems to cry out for unification as an enforcement strategy. We are not concerned with performance issues at this stage; neither do we intend to explore all avenues of declarative programming. For historical reasons, we concentrate on a Prolog-derivative for our logic meta language. Finally, we make the symbiosis between the two paradigms explicit by allowing base-level programs to be expressed as terms, facts or rules in the meta-level; we will refer to this as a *representational mapping*.

```
class Array {
  private int[] contents;

  Array(int sz) {
    contents = new int[sz];
  }

  int getAt(int i) {
    return contents[i];
  }

  void setAt(int i, int e) {
    contents[i] = e;
  }
}
```

Figure 3a: a Java Array

```
Class(Array<?E1>, {
  private ?E1[] contents;

  Array<?E1>(int sz) {
    contents = new ?E1[sz];
  }

  ?E1 getAt(int i) {
    return contents[i];
  }

  void setAt(int i, ?E1 e) {
    contents[i] = e;
  }
}
```

Figure 3b: a generic Java Array

Consider the simple Java class in figure 3a: it implements an array of integers. Next to it in 3b the original class has been embedded in a meta-declaration using a representational mapping. The notation is fairly crude and to clarify it somewhat elements from the meta-program have been highlighted. Notice that the original element type was replaced by a logic variable.

This is an example of using LMP for code-generation; it was explored in [KDV98] under the exotic name *TyRuBa*. A proper query substituting `int` for `?E1` in 3b would produce 3a.

A totally different way to view LMP is introduced in [RW98] as the *Smalltalk Open Unification Language* (SOUL). This approach actually applies constraints specified at the meta level to the base level program. The representational mapping is based on the presence of predicates that give access to syntactic elements belonging to the base level.

```

uses(?c1, ?c2, ?p) :-
    class(?c1),
    method(?c1, ?m),
    calls(?m, ?c3),
    (?c3=?c2 | not(member(?c3, ?p)),
        uses(?c3, ?c2, [?c3 | ?p])).

recursive(?c) :-
    uses(?c, ?c, []).

```

Figure 4: a recursion test

In the above example we assume the availability of predicates `class`, `method` and `calls` to access the structure of a base program. The `uses` rule establishes a transitive calling path between two classes, which allows us to derive the `recursive` rule. This in turn could be used to enforce the acyclicity constraint from the manager-employee example.

4. LMP AND ASPECT ORIENTED PROGRAMMING

In [KDV98] a LMP framework is proposed to support sophisticated type systems for statically typed programming languages such as Java. This framework, called *TyRuBa*, turns a type system into a computationally complete environment and allows a programmer to specify the static structure of a program as a set of logical propositions. In one of the next sections we will report on an experiment to use *TyRuBa* as a system to describe software architectures with. In this section we will build on the relationship between LMP and *Aspect Oriented Programming* (or AOP for short). We refer to [DVD99] for an extended bibliography; suffice it to say that AOP is concerned with the production of software as a result of a *weaving* process. The weaver is an AOP-related tool that is capable of merging aspects of a software application, each of them described in a specific aspect language.

In [DVD99] it is proposed that a LMP may well function as an aspect-oriented programming environment. As evidence for this, a well-known case for AOP (synchronisation of cooperating processes using an aspect language called COOL) is expressed in *TyRuBa*. An important conclusion from this experiment is the fact that a general purpose framework, in casu LMP, can be used to host aspect programs; hitherto, aspect languages were specific to the aspect under consideration.

In deference to the subject of this contribution, we will not concentrate on technical applications of AOP; instead we will consider an interesting application of AOP involving design as much as implementation. In [DD99] the idea is launched that *domain knowledge* might well constitute an aspect in the AOP sense. Separating some problem into its domain aspect and its implementation aspect by describing them in a some aspect language and then

producing a piece of software by applying a weaver seems a very attractive approach. Moreover, it seems to fit very well with the concept of software co-evolution introduced earlier on.

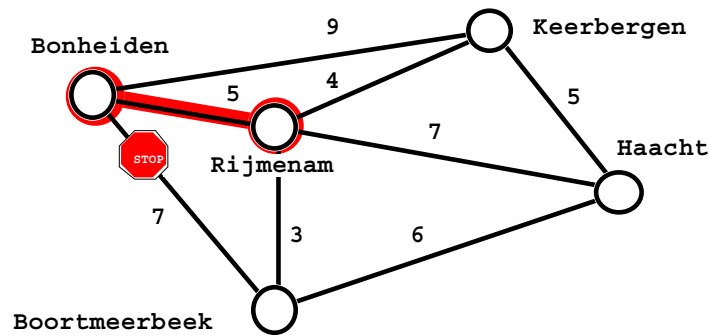


Figure 5: a shortest path problem

Figure 5 represents the test case proposed in [DD99] to explore this idea. The example is taken from a GIS-application, involving a mix of a conventional algorithm to compute a shortest path, and the specifics of the domain, which allow us to improve the basic algorithm.

```

branchAndBoundFrom: start to: stop
| bound |
bound := 999999999.
self traverseBlock:
[:node :sum |
node free ifTrue:
[sum < bound ifTrue:
[node = stop
ifTrue: [bound := sum]
ifFalse: [self branch: node sum: sum]]].
self traverseBlock value: start value: 0.
^bound

```

Figure 6: the branch-and-bound program

In order to keep things as simple as possible, we consider an elementary *branch-and-bound* strategy. In figure 6 this is implemented using an auxiliary `branch:` method in order to fix the order in which branches are selected.

```

branch: node sum: sum

node free: false.
node edges do:
[:edge | self traverseBlock value: edge next
value: sum + edge distance].
node free: true

```

Figure 7: fixing the selection order

Figure 7 contains a possible implementation for `branch`: and it contains an enumeration of all possible edges leaving a node. However, the message `edges` is no longer resolved by the base program, but by a query in the logic meta program containing the knowledge about this particular domain:

```

Fact city (Rijmenam)
Fact city (Boortmeerbeek)
...
Fact road (city (Rijmenam),city (Boortmeerbeek), [3])
Fact road (city (Keerbergen),city (Rijmenam), [4])
...
Fact prohibitedManoeuvre (city (Rijmenam),city (Bonheiden))
Rule roads (?current,?newResult)if
    Findall (road (?current,?next,?distance),
            road (?current,?next,?distance),?result)
    privateRoads (?current,?result,?newResult)
Rule privateRoads (?current,?result,?newResult)if
    prohibitedManoeuvre (?current,?next),
    removeRoad (?result,road
(?current,?next,?distance),?newResult)
    Fact privateRoads (?current,?result,?result)

```

The particular flavour of LMP we use here is the *Smalltalk Open Unification Language* (or SOUL) introduced in [RW98]. The query would look something like this:

```

... Roads (city(Rijmenam), ?result)

```

In [DDW99] an explanation is given of how the base program and the meta program communicate. The basic idea is to effect a kind of linguistic symbiosis (see e.g. [PS94]) based on a two-way reification of language entities; in the case of SOUL this amounts to wrapping Smalltalk objects inside Prolog facts and vice versa.

5. LMP AND SOFTWARE ARCHITECTURES

In [MWD99] LMP is explored as a framework in which to express software architectures. SOUL is proposed as both the target as the medium for this study: it is used as a kind of architectural description language and it is applied to the architecture of SOUL itself.

The kernel of SOUL is a logic query interpreter with the following architecture:

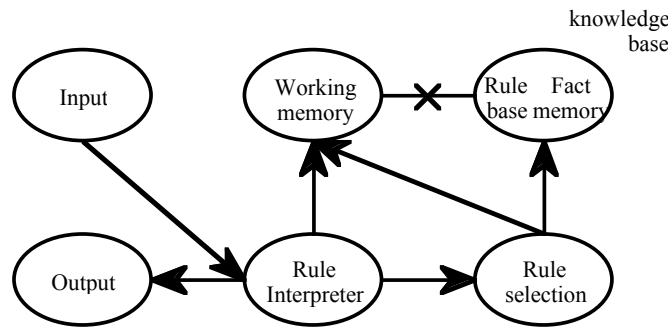


Figure 8: the SOUL rule base architecture

This architecture is representative for rule bases in general; moreover it is sufficiently challenging to be used as a case study.

6. *LMP* AND SOFTWARE COMPONENTS

Xxxx

7. CONCLUSION

xxxx

8. BIBLIOGRAPHY

[CL97] Documenting Reuse and Evolution with Reuse Contracts

Carine Lucas

PhD dissertation, Vrije Universiteit Brussel (1997)

[DD99] Is domain knowledge an aspect?

Maja D'Hondt and Theo D'Hondt

Proceedings of the ECOOP99 Aspect Oriented Programming Workshop (1999)

[DDW99] Using Reflective Programming to Describe Domain Knowledge as an Aspect

Maja D'Hondt, Wolfgang De Meuter and Roel Wuyts

Proceedings of GCSE '99 (1999)

[DVD99] Aspect-Oriented Logic Meta Programming

Kris De Volder and Theo D'Hondt

Proceedings of Reflection '99 (1999)

[JB99] Syntactic Abstractions for Logic Meta Programs, or vice-versa

Johan Brichau

Draft publication (1999)

[KDH98] A Novel Approach to Architectural Recovery in Evolving Object-Oriented Systems

Koen De Hondt

PhD dissertation, Vrije Universiteit Brussel (1998)

[KDV98] Type-Oriented Logic Meta Programming

