Chapter 5

An overview of Mirjam and WeaveC

Authors: István Nagy, Remco van Engelen, Durk van der Ploeg

Abstract In this chapter, we elaborate on the design of an industrial-strength aspect-oriented programming language and weaver for large-scale software development. First, we present an analysis on the requirements of a general purpose aspect-oriented language that can handle crosscutting concerns in ASML software. We also outline a strategy on working with aspects in large-scale software development processes. In our design, we both re-use existing aspect-oriented language abstractions and propose new ones to address the issues that we identified in our analysis. The quality of the code ensured by the realized language and weaver has a positive impact both on maintenance effort and lead-time in the first line software development process. As evidence, we present a short evaluation of the language and weaver as applied today in the software development process of ASML.

5.1 Introduction

One of the primary goals of the Ideals project is to develop methods and tools to improve the handling of crosscutting concerns, such as tracing and profiling. As a solution, a proof-of-concept aspect-oriented [38] language for the C language [64] and weaver (called WeaveC1) have been proposed and developed by Durr et al. [33]. Subsequently, a case study has been carried out on a representative component of the ASML software to assess the usability of WeaveC for four particular crosscutting concerns. The outcome of the case-study has shown that these four crosscutting concerns are

---

1Here, the name WeaveC refers both to the AOP language and weaver.
manifested in a significant amount of code (20-40%) besides the code representing the original concerns of the component. By using the CoCoMo model [5], the case-study estimated 5-10% effort reduction and 2-3% lead-time reduction for the code developed with WeaveC, as compared to a non aspect-oriented solution in the C programming language. Although these numbers may seem to be small, it is important to note that all software development (approximately 500 software developers) is affected by the above mentioned and other crosscutting concerns.

Briefly, the case-study demonstrated the industrial maturity of aspect-oriented programming by means of a successful proof-of-concept. As a consequent step, a transfer project was initiated by the industrial partner with the following goals:

- Develop a general purpose aspect-oriented language for the C language that is capable of modularizing crosscutting concerns within ASML software.
- Implement a robust, industrial strength weaver for the proposed aspect-oriented language.
- Develop a way of working (i.e. micro-process) that describes the necessary developer roles, activities and artifacts to deal with crosscutting concerns (by means of the above described programming language and weaver).

The remainder of the chapter is organized as follows: Section 5.2 presents an analysis to identify the requirements of a general purpose AOP language. We also outline a strategy on working with aspects in large-scale software development processes. Section 5.3 discusses the concepts of the proposed aspect-oriented language, weaver and micro-process. Section 5.4 presents the results of an evaluation on the proposed language and weaver. Finally, Section 5.5 draws conclusions.

5.2 Problem analysis

In this section, we present an analysis to identify the requirements towards a general purpose aspect-oriented language to handle crosscutting concerns in the ASML context. The term ‘ASML context’ covers two important design considerations: (a) the existing solution designs (i.e., idioms, see Section 2.2.1 in Chapter 2) of crosscutting concerns and (b) the way of working of software developers within ASML. In the following subsections, we elaborate further on these considerations.

5.2.1 Boundaries of modularization

An important concern of the analysis is to explore the boundaries of modularization of the current solution design of crosscutting concerns. The aim of this step is to determine the set of necessary AOP language abstractions and variation points required

\[2\text{The proof-of-concept AOP language had only a limited set of language abstractions that were necessary for the execution of the case-study.}\]
for the proper modularization. Our objective is not to come up with new language
abstractions; in contrast, our objective is to re-use the existing language concepts of
aspect-oriented languages as much as possible. For this purpose, we iterate over the
essential concepts of a reference model of AOP languages [79, 7]:

1) static ZDSPTI_module_handle ti_handle ; // (2)
2) ...
3) int ZDAPSF_startup ( int sim_mode, boolean caching)
4) {
5)   int result = 0 ;
6) ...
7)   THXATrace("KI", (1), "ZDAPSF_startup", "> (" // (1)
8)     "sim_mode = %i, caching = %b"
9)     ")"
10)    , mode, caching // (1)
11) }
12) ...
13) ZDSPTI_timing_handle timing_hdl = NULL ; // (3)
14) if ( result == OK ) // (2)
15)   result = ZDSPTI_register_module("ZDAPSF", &ti_handle); // (2)
16) ...
17)   ZDSPTI_timing_in(ZDSPTI_func_timing_hdl, ti_handle, // (3)
18)     "ZDAPSF_startup", &timing_hdl); // (3)
19) ...
20)   if ( result == 0 )
21)   {... }
22) }

Listing 5.1. Illustration of three particular concern instances - Tracing, Setup
of Timing and Start of Timing - as they may\textsuperscript{a} appear in a particular C
function in ASML software. Respectively, these concerns are represented by
the lines marked (1), (2) and (3).

\textsuperscript{a}The concern instances presented here are slightly altered for reasons of confidentiality. How-
ever, this does not affect the essence of the examples.

\textbf{Base Language} Aspect-oriented languages are considered to be extensions of generic
programming languages. The fact that ASML software is a legacy system written in
the C programming language, restricts the base language to C. Note that the choice of
base language will restrict further the design space of other aspect-oriented language
abstractions (e.g., types of join points).

\textbf{Join points} In AOP, (behavioral) join point corresponds to events in the control flow
of a program. The example of Listing 5.1 shows that the concern instances typically
appear at the beginning of the execution of a function. In other words, the example
concern instances need to be executed before (and also after) the occurrence of the
event ‘execution of a function’. Considering the current idioms that realize the crosscutting concerns, we identified two necessary types of join point: call and execution of a function.

**Context of Join points**  Listing 5.1 shows that various types of context information are required for the modularization of a crosscutting concern. For instance, the concern instance of Tracing refers to the name of the function in Line 7, and formal parameters of the function in Line 10. The concern instance of Setup of Timing refers to a local variable of the function (‘result’) in Line 14. The same concern instance refers to the name of the module of the function (‘ZDASP’) in Line 15. Besides these types of context information, we identified that the following types of context information are required: global parameters of a function, types of parameters and local variables, return type of functions and the fact whether a parameter or local variable behaves as an input and/or output parameter in the function (derived information from data-flow analysis).

**Context and join points related to pointcuts**  AOP languages generally make use of the above described context information, not only in the parametrization of crosscutting behavior but also in the designation process of join points (i.e. in pointcut or query languages of AOP languages). Obviously, we will also use these in the designation process of join points.

**Context Originated from Build Process**  In our investigation, we recognized that information derived from the build process is also used in the idioms that realize crosscutting concerns in software. For instance, the concern instance Tracing refers to the component code ("KI") in Line 7; this information is determined from the target of the build process. In a broader view, when crosscutting concerns are woven into different products of a product line, product (and platform) specific information also needs to be addressed in the modularization. Hence, information about product and platform can serve as variation points; these are typically originated from the configuration of the build processes.

**Advices and Variation Points in Advices**  Advices are the units of AOP languages to formulate the crosscutting behavior in terms of the instructions of generic programming languages. All the earlier described types of context information — i.e., the join point, the properties and relationships of a join point, and the platform and product specific information from the build process — can serve as variation points in the formulation of an advice. Normally, these variation points can be directly used/referred to through pseudo variables (e.g. thisJoinPoint in [65]) and/or through parameters that are extracted from pointcuts and passed to advice bodies (e.g. context exposure in AspectJ).

Besides the direct references and usage, we recognized indirect usages of the above listed types of context information. For instance, the string literal ‘mode = %i,
caching = %b’ in Line 8 contains the name of formal parameters and format specifier characters based on the type of the formal parameters. That is, the formal parameters (as join point context) are not explicitly referred to as variations points (unlike in Line 10) but their properties need to be used to compute a variation point (cf. the string literal in Line 8) in the notation of the advice-concept.

The fact that we need to deal with a legacy system may have further constraints on the design space of the advice-concept: the crosscutting behavior should be formulated in terms of instructions that can express calls to existing software libraries. For instance, the calls THXATrace() in Lines 7-10 and ZDSPTI_timing_in() in Line 17 are such ‘legacy’ calls to libraries that realize tracing and timing.

A consequence of this constraint is that the concept of variable-length argument list - denoted by ‘...' in C - needs to be treated as variation points in the notation of an advice. For instance, in Listing 5.1, the concern instance of tracing is represented by the call THXATrace("KI", (1), "ZDAPSF_startup", "> (" "sim_mode = %i, caching = %b" ")", sim_mode, caching) with 7 arguments, in Lines 7-10. The arguments of this call statement depends the formal parameters mode and caching of the function int ZDAPSF_startup ( int sim_mode, boolean caching) This means that the concern instance of tracing in a different function context will be represented by a call statement with different format string and different number of arguments. For instance, the function ZDAPSF_shutdown() without any parameter will have the tracing call THXATrace("KI", (1), "ZDAPSF_startup", "> (" ")") with only 5 arguments. To address this problem, the notation of the advice concept needs to able to deal with the concept variable-length argument list as a variation point. This variable-length argument list can always be derived from the actual join point and its context (cf. the base function to be woven and it formal parameters).

Note that re-implementing the tracing library with a suitable interface is not an option either, as it would introduce other maintenance problems related to legacy systems.

Aspectual States/Aspect Instances In Listing 5.1, both the concern instances of Setup of Timing and Start of Timing use the variable ti_handle that represents the module handler of Timing. The concern instance of Timing uses another variable, called timing_hdl in Line 17. The difference between these two variables is apparent from the example already: ti_handle is declared as a global static variable in Line 1. This means that ti_handle can (and will) be used in every function of the module; thus, it can be shared among different concern instances that appear in different functions. In contrast, the variable timing_hdl is declared as a local variable of the function in Line 13; hence it is local only to those concern instances that appear only within the same functions. In terms of aspect-oriented programming, we identified the need of two types of aspectual states: one which is shared only among those advices that are woven at the same join point (per join point) and one which is globally shared among every advice (i.e., per aspect).

Note that there are many other fundamental concepts, e.g. ordering of advices, context exposure, et cetera, in the design space of aspect-oriented languages. Naturally,
we consider these concepts as parts of our design space; however, the discussion on the motivation for each particular language concept is beyond the scope of this chapter.

5.2.2 Quality aspects of concrete syntax for large-scale development

In the previous section, we discussed the boundaries of modularization that determines the necessary expressiveness of the language - i.e., abstract-syntax of a language, in terminology of Domain Specific Languages [42]. However, besides necessary expressiveness, there are various quality aspects on the concrete syntax that are significant from the point of view of large-scale software development. In this section, we discuss these.

**Predictability (in design phase)** Predictability ensures developers that certain properties are held during the development of software. This is crucial in large-scale development to prevent mistakes and errors already in the design phase as soon as possible. Besides, when introducing new technology into a large organization, ensuring predictability in the design phase is well-motivated for the following two reasons: (1) it can lessen the risk of improper use (and its undesired side effect) of the new technology; (2) it can reduce the fear of using a new technology among the developers (e.g., by providing well-controlled means for the new technology).

To this aim, we are going to use well-known language mechanisms, such as type-checking, enforcement of declarative completeness, and an extensive set of syntactic and semantic rules in the design of the aspect-oriented language.

**Evolvability** Evolvability is an important software quality factor that indicates the ability of developing programs incrementally. In general, evolvability facilitates extending an application towards new requirements mostly by reusing previously written modules with minimal or no modification. By ensuring evolvable specification, we expect to reduce the complexity of the code caused by the frequent appearance of the phenomenon called ‘deviations from standard functionality’ in a large-scale development.

To this aim, we are going to provide language abstractions that (1) can be re-used in different specifications and (2) can support the re-use of existing specifications.

**Extendibility** In principle, language constructs that provide means of parametrization can positively contribute to the extendibility of a language. Providing ease of extendibility for a language is beneficial in large-scale development for two reasons. First, due to time-to-market pressure the language (and also the weaver) needs to be incrementally delivered. Secondly, rolling out a new version of a language and weaver is not a trivial task as it affects many ongoing software development activities; hence, a

---

3These language mechanisms obviously require adequate compiler and/or run-time support.
large number of development modules maintain dependencies to aspect specifications. This means that change requests towards the language should result in changes in the notation as *minimal* and *isolated* as possible.

**Comprehensibility**  We define comprehensibility as the ability to understand the meaning of a program by just looking at its source code. Comprehensibility can be influenced by the programming style and the notation of the abstractions of the adopted language, such as how the program units are modularized, where the references in the units are specified, and the style of notation that reflect the underlying computational paradigm. Although this quality aspect had less significance compared to the previous ones in our list, we needed to take into account the fact that developers at ASML have a stronger background on procedural and object-oriented languages (with strong typing) compared to logic or functional languages.

### 5.2.3 Expected way of working

In the previous two sections, we discussed the requirements towards an aspect specification language used in large-scale software development. However, when introducing aspect-oriented programming (or any other new technology) into a large organization, a clear and sound strategy on the expected way of working is also necessary, besides the required means and artifacts. In this section, we discuss our strategy on how aspects are intended to be used in the software development process of ASML.

To make sure that developers can benefit from the advantages of aspect-oriented programming, the objective is that *every* developer should be able to *use* aspects easily with a minimal learning curve. On the other hand, to minimize the possible danger of the improper use of AOP, only *a few* of the developers are allowed to *write* and *release* new aspects. To this aim, our objective is to provide generic aspects with highly and easily customizable interfaces:

- Developers are expected to use aspects in a standard way (seamlessly, by enabling them in the build process).
- Most users will use aspects with their standard functionality.
- Some users will want to deviate from the standard functionality. In other words, they want to customize the functionality of generic aspects according to their special needs. For instance, some users will want to ‘switch off’ the tracing of certain time-critical functions.
- This customization should be minimal and rely on design information added to the source code, as most users are not allowed to modify aspects or write their own one.

---

4 E.g., locality of changes, consistency and clarity of code, et cetera.
5 E.g., undesired side-effects in the control-flow, ‘patching by aspects’, et cetera.
5.3 Mirjam, an aspect-oriented language extension for C

In this section, we outline the important constructs and characteristic properties of the realized aspect-oriented language, called Mirjam. Due to a lack of space, we cannot iterate over the full design space of Mirjam. The interested readers can find a more detailed introduction and description of the language in [9].

5.3.1 Aspect

The main unit of modularization is the aspect specification file. An aspect specification may contain two language constructs (see Listing 5.2): context declaration closure and aspect declaration.

```c
1) context{
2)     #include "THXAttrace.h"
3)     #define FALSE 0
4) 
5)     typedef int boolean;
6) }
7) 
8) aspect SimpleAspect
9) {
10)     advice someAdvice() before (FunctionJP JP)
11)     {
12)         boolean tracing_flag = FALSE;
13)         THXAttrace(JP.module.name, JP.name, tracing_flag, "> ");
14)     }
15)     ...
16) }
```

Listing 5.2. Illustration of the context declaration closure and aspect declaration in an aspect specification file.

The context declaration closure is a placeholder for a standard C declaration. We will use the declarations in C in the context closure to ensure declarative completeness in the notation of advices. The context closure in the listing above has two preprocessor directives in Lines 2 and 3, and a typedef declaration in Line 5. To resolve the preprocessor directives, the aspect specification is first preprocessed by a standard C preprocessor. The declarations in the context closure are used later in the specification of the crosscutting behavior (i.e., advice in term of Mirjam) in Lines 11 and 12. Note that the function call ‘THXAttrace’ (in Line 13) is declared in the included THXAttrace.h file.

The aspect declaration part is a container type of program element: it works as a name space. The contained language abstractions (query, advice- and binding-declarations) can be referred to through this name space.
5.3.2 Queries as pointcuts

The language concept of pointcut is realized by the language abstraction query in Mirjam. A query in Mirjam returns a set of join points. A join point is a tuple that contains an arbitrary number of (but at least one) tuple elements. A tuple element can be of two types: join point location and join point context. As the names suggest, the type of join point location describes the place where we want to weave in crosscutting behavior (e.g., execution of a function), while the type of join point context describes the context of the weaving. We can use the context information for (1) either refining of the weaving location or (2) customizing the crosscutting behavior to be woven to certain weave contexts.

```
1) int a;
2) int f(int b, int c) {
3)    ...;
4) }
5) int g(double d) {
6) }
7) }
```

Listing 5.3. An example of base code that we will use in the rest of the discussion to illustrate how the query language of Mirjam works.

We will use a number of examples of queries to present the characteristic features of the query language of Mirjam:

```
1) query Q1() provides (FunctionJP JP) {
2)     JP: true;
3) }
4) }
5) query Q2() provides (FunctionJP JP) {
6)     JP: JP.name =~ "f.*";
7) }
8) }
9) query Q3() provides (FunctionJP JP, Variable@JP V) {
10)    JP: true;
11)    V : true;
12) }
13) }
14) query Q4() provides (FunctionJP JP, Variable@JP int V) {
15)    JP: tuple(JP) in Q1() && |JP.formalParameters()| >= 1;
16)    V : V in JP.formalParameters();
17) }
18) }
19) }
20) }
21) }
```
22) query Q5() provides (FunctionJP JP, Variable@JP[] V)
23) {
24)    JP: tuple(JP) in Q1() && |JP.formalParameters()| >= 1;
25)    V : V == JP.formalParameters() ;
26) }
Listing 5.4. Examples of query-declarations in Mirjam.

As an example of a query, consider the query declared between Lines 1 and 4 in Listing 5.4. The query declaration starts with the keyword `query` followed by an identifier (`'Q1'`) and a list of possible formal parameters. After the formal parameters we define a set of tuple variables preceded by the keyword `provides`. The list of tuple variables describes the elements and type of tuples that the given query provides. In Lines 1-4, the query Q1 provides tuples with one tuple element, the type of this tuple element is `FunctionJP`. The type `FunctionJP` represents the type of the join point location of the execution of a function. Inside the query, a tuple condition needs to be defined for each tuple variable. In Line 3, the tuple condition ‘true’ means that every particular function execution will satisfy this query; i.e., there is no further restriction on the tuple variable `JP`. The resulting set of tuples of query Q1 executed on the base code of Listing 5.3 is \{(f), (g)\}.

In Lines 6-9, the query Q2 is similar to Q1 except that it has a more restrictive tuple condition in Line 8: the name of the executing function shall start with the prefix ‘f’ (defined by the regular expression ‘f.*’). Similar to the notation of object-oriented programs, the dot operator can be used to access instance variables and execute methods of certain types in Mirjam. The result set of Q2 on the given base code is (f), as there is only one function that can satisfy this condition.

In Lines 11-15, the query Q3 will provide set of tuples with two tuple elements. The first tuple variable is of type ‘function execution’. The second, new tuple variable is of type `Variable`, which is a type of join point context. More precisely, the type `Variable` can represent global variables, formal parameters and local variables of an executing function. In Line 11, the declaration `Variable@JP` means that we expect the variables in the scope of the function execution of JP; that is, there is direct link between a particular function and variable in the result set of the query. Note that a second tuple condition ‘true’ has been declared for the second tuple variable in Line 14. When a query is evaluated, the query engine iterates over the possible values of both tuple variable types. During the iteration, if there is a combination of particular values of tuple variables that can satisfy both tuple conditions (based on the available base code), a tuple is created and added to the result set of query. In short, the query Q3 returns an ordered set of tuples (i.e. the Descartes-product) of all functions and variables related with those functions: \{((f),a), (f,b), (f,c), (g,a),(g,d)\} based on the code of Listing 5.3.

In Lines 17-21, the query Q4 illustrates a bit more complex use of queries. The tuple variable `V` in Line 17 is declared with a type restriction: the C type of the variable

6For the sake of simplicity, the letters f and g represent the execution of functions f and g.
is restricted to int. In Line 19, we show a possible re-use of a previously declared query: the tuple variable JP is converted to a tuple by the conversion operator ‘tuple’. The membership relation ‘in’ specifies that JP (as converted to a tuple of one element) should be in the result set of the query Q1. The membership relation (‘in’) is not the only way to re-use queries; the equivalence relations can also be used with query calls. Queries can be called with formal parameters to perform selection or projection (in terms of tuple relational calculus [24][7] on previously defined queries. In Line 19, the second part of the tuple condition specifies that the executing function needs to have at least one formal parameter. The second tuple condition in Line 20 specifies that V is a formal parameter of JP. The execution of Q4 on the provided base code of Listing 5.3 results in the following set of tuples: { (f,b), (f,c) }.

In Lines 23-27, the query Q5 is a slightly modified version of the query Q4. Q5 differs from Q4 in two places. First, the type of the tuple variable V is an array type denoted by the symbols ‘[]’ in Line 23. Note that this tuple variable has no C type specification either. Secondly, the tuple condition uses the equivalence relationship (‘==’) instead of the membership relationship (‘in’). This means that Q5 will provide tuples with two elements: a tuple element of a function-execution and a tuple element of an array of variables. The equivalence relationship indicates that the array of variables must be the array of formal parameters of the corresponding function. This means the result set of Q5 is { (f, {b,c}), (g, {d}) } based on the code of Listing 5.3.

### 5.3.3 Advices

Similar to other AOP languages, the program unit that specifies the crosscutting behavior is called advice in Mirjam. The crosscutting behavior is defined in terms of instructions of the C programming language.

```c
1)  advice printIntParam(Variable@JP int V) before {FunctionJP JP)
2)  {
3)   printf("In module %s, function %s executes with argument %s=%d",
4)       JP.module.name, JP.name, V.name, V);
5)  }
```

Listing 5.5. An example of advice-declaration.

As an example of an advice, consider the advice declared in Listing 5.5. The declaration starts with the keyword ‘advice’ and is followed by an identifier and a (possibly empty) list of formal parameters. A formal parameter has to have a type specifier in Mirjam and may have a type specifier in C. The formal parameters are followed by the type of the advice that specifies whether the advice is to be executed before or after the execution of a join point. Finally, the last element in the signature of the advice is the actual join point variable in the form of a formal parameter (FunctionJP JP). In the

---

[7] Defining the formal semantics of queries is beyond the scope of this chapter.
body of the advice, we refer to the properties of the join point through this variable in Line 4. In the same line, we also use the parameter V as a formal parameter similar to the formal parameters used in C.

Note that the correct use of the format specifier (e.g., %d in Line 3) cannot be checked just like in normal C code. For instance, we could use %f that would have resulted in an incorrect execution. This piece of advice exemplifies another possible problem as well: if we coupled it with the query Q4 (of Listing 5.4) the advice would be woven two times on the function f, as Q4 returns {(f,b), (f,c)}. This means that the execution of the program would have two output messages, see Listing 5.6.

6) In module ZZ, function f executes with argument b=5
7) In module ZZ, function f executes with argument c=11

Listing 5.6. Illustration of output messages from advice executions.

This execution is correct because Q4 exposes only formal parameters of type int. However, if the query is not restrictive enough, formal parameters with other types than int (e.g. double, pointer-type, et cetera) can be queried too. This would result in incorrect output messages, since the advice expects a parameter of type int. In addition, having the same message two times, with only a small difference at the end, is not considered a neat solution: a single printf statement could handle a variable number of arguments. So the question is how to specify the statement printf with variable number of arguments of different types that depend on the weaving context? Note that we discussed the same problem in Section 5.2.1 at ‘Variation Points in Advices’.

5.3.4 Advice generators

To handle the category of problems described above, we introduced the language construct advice generator. These are special built-in constructions that can be used in the body of an advice and are recognizable by the form @type<>@. The first element of a generator is usually an iterator of the form [iterator: expression]. The iterator contains the name of an array variable and the (local) name for each element, and returns a list of expressions. The expression will be evaluated for each element. Finally, the type says how the expressions are combined and what the outcome of the advice generator is.

---

8 To prevent this sort of errors, we do type checking between the type of the actual tuple values and the type of the advice-parameter.
9 In other (typically, functional) languages one can achieve the same functionality by using the list manipulator functions map and join on lists.
For an example application of advice generators, consider Listing 5.7. In Lines 4-8, the advice generator `@StringConstant` will generate a format string for an array of variables passed to the advice (Args). In Line 6, the built-in function `strConcat` creates a piece of string from the name of a variable (cf. formal parameter of a function), from the equation sign (`=`) and from the result of the function `formatString` returning the format specifier of the C type of the variable. In Line 5, this concatenation function is applied on each variable in the array of the variables ('arg in Args'). This will result in a list of a piece of format string for each formal parameter. The resulting list of format strings per parameter will be combined into a final string literal by the StringConstant generator, using the semicolon as a separator. Similarly, a variable-length argument list will be created by the advice generator `@VarargExpression` in Line 10. If we couple this advice (Listing 5.7) with the query Q5 (in Listing 5.4) the advice is woven once on the functions f and g, as Q5 returns (f, {b,c}), (g, {d}). This means that the execution of the program would have two output messages, see Listing 5.8.

Listing 5.8. Illustration of output messages from advice executions with advice generators.

5.3.5 Bindings

The third language construct that can be declared within an aspect is the binding closure. The binding closure contains typically one or more binding definitions introduced...
by the keyword *foreach*. One binding definition can couple a query with one or more *binding entities*. Different types of binding entities exist in Mirjam: for example, the binding entity *apply* binds advices to queries, and passes the result of the queries to the advices using a parameter-passing mechanism. When more than one advices are to be applied to the same join point, one can specify the order of their application by the binding entity *order* per join point. The binding entity *error* and *warn* can raise errors, when certain situations occur in the base code. Finally, the binding closure and binding entities may also contain variable declarations, called binding variables, to share states between advices. In the following subsections, we will show the use of each of these language constructs.

1) aspect VerySimpleTracing
2) {
3) 
4) query Q5() provides (FunctionJP JP, Variable@JP[] V)
5) {...}
6) 
7) advice printParams(Variable@JP[] Args)before(FunctionJP JP)
8) {...}
9) 
10) binding
11) {
12) foreach (thisFunc, vars) in Q5
13) {
14) apply on thisFunc {
15) printParams(vars);
16) }
17) }
18) }

Listing 5.9. An example of binding and an illustration of passing context information to advices.

For an example of a binding-declaration, consider Listing 5.9. In Line 3, the query Q5 provides a set of tuples of a join point variable (FunctionJP JP) and a context variable (Variable@JP V). In Line 11, the binding definition iterates over the result set of Q5: *thisFunc* and *vars* will represent the two variables of each tuple. For each iteration step, these variables hold the values of the actual tuple elements. In Lines 13 and 14, the actual values of *thisFunc* and *vars* are passed as parameters to the advice in the advice-call. Finally, in Line 6, the variables are ‘received’ and they are used in the body of the advice. This type of parameter passing works in a similar way to the macro substitution mechanism of the C preprocessor.
5.3.6 Bindings variables

As we mentioned in the previous section, the binding closure and binding entities may also contain variable declarations, called binding variables, to share states between the executions of advices.

```
1) binding {
2)     static int ti_handle;
3)     foreach (startup) in startupFunction()
4)         apply on startup {
5)             moduleRegistration(ti_handle);
6)         }
7) }
8) }
9) }
10) foreach (f) in timedExecution()
11) {
12)     apply on f {
13)         int timer;
14)         functionTimingStart(ti_handle, timer);
15)         functionTimingStop(timer);
16)     }
17) }
18) }
19) }
20) }
```

Listing 5.10. Illustration of the use of binding variables.

Consider the variables static int ti_handle and int timer in the Lines 2 and 14 of Listing 5.10. One can use binding variables per aspect (like ti_handle in Line 2) and per join point (like timer in Line 13). ‘Per aspect’ binding variables, e.g., ti_handle, are shared between all advices within a binding. This means that if one advice changes ti_handle, the change can be observed in all other advices of the aspect. ‘Per join point’ binding variables, e.g. timer, are shared only among those advices that are called from the binding entity where the variable is declared. For example, timer is shared only between the execution of the advices functionTimingStart and functionTimingStop in each function where they are applied together.

5.3.7 Annotations

Annotations in WeaveC are used to attach semantic meaning to syntactical constructs in a program written in the programming language C. The annotation mechanism is designed such that the following objectives are met:
Easy of use An application of an annotation to a program should have minimal overhead, both in terms of the effort spent by a developer and of the impact on the program listing. The syntax should be unobtrusive (not disturb the natural flow of the program listing) and similar to the syntax of similar C concepts (type modifiers like storage class, const et cetera).

Flexible for different usage The application of annotations is only limited by the scope of tools that can process them, so the mechanism should place few restrictions and allow for varied use.

Allow checking of the correct application of annotations This means that it must be detectable by a tool if an annotation is applied at the wrong location (i.e., to a function while it makes no sense for a function), with the wrong arguments (i.e., with a value for a specific field while the annotation has no such field) or multiple times (when it makes no sense to do so for the annotation). It does not need to be possible to check for unexpected absence of annotations (i.e., a function has no annotation while one is always expected) or illegal combinations of annotations or values of fields of annotations (i.e. a function can not have both annotations ‘a’ and ‘not_a’). These checks, if desired, must be performed by the tools that make use of the annotations.

Easy to hide from a program This should be the case for instance if the program is to be read by tools that are unaware of this annotation mechanism. Such hiding should be possible by a standard-compliant C pre-processor or a standard text processing facility.

Robust Common typing errors in the declaration or application of an annotation should not lead to undesirable parsing of the actual program code.

5.3.8 An example of using annotations

In this section, we outline of how the annotation mechanism of WeaveC can be used.

```
1) // --- an annotation-declaration in .c files ---
2) /*$
3)   annotation{
4)       boolean value;
5)   } trace[module, function, parameter, variable, type] = {
6)       value = TRUE
7)   };
8) */$
9) 
10) // --- an annotation-application in .c files ---
11) int critical_function()
12) /*$trace(FALSE)$*/
13) {...}
```
Listing 5.11. Illustration of the declaration, application and use of annotations.

The declaration and application of annotations are denoted by the tokens /*$ and $*/, in the style of Splint [40]. In Listing 5.11, the annotation trace is declared between Lines 2-8. The annotation declaration has one boolean field called value (in Line 4) and it can be applied on modules, functions, formal parameters, variable declarations and type declarations, as indicated in Line 5. In Line 6, we specify that the default value for the field value is TRUE\[11\]. An application of the annotation trace is illustrated in Line 12: the application is part of the signature of the function declaration int critical_function(). Finally, the query allTraceableFuncs() shows a use of annotations for determining join points in Mirjam, in Lines 18-19. Only those function executions will be designated that do not have the name "main" and do not have the application of the annotation ‘trace’ with the value FALSE.

5.4 Evaluation

Dürr et al. [34] carried out an experiment on quantifying the benefits of using aspect-oriented programming by means of the above defined language and weaver. The setup of the complete experiment and all statistical data about the results can be found in [34]. In this section, we provide only a summary about the setup and the results of the experiment.

The goal of the experiment was to determine whether using AOP speeds up the development and maintenance of the ASML codebase. The experiment consisted of five change scenarios in two sets related to use of the concern Tracing. Twenty software developers participated in the experiment\[12\]; the participants were split into two groups. The most important reason to do the splitting was to verify that the two sets of change scenarios were equivalent.

Although it is hard to extract significant results from the experiment due to small number of participants, the following conclusions can be definitely supported from the results of the experiment:

- Adding tracing to a function takes considerably less time with AOP than without.

\[11\] This is a standard predefined macro. The weaver has a built-in C preprocessor to resolve macro substitutions within annotation declarations and applications.

\[12\] The experiment was actually combined with training on the use of aspects and WeaveC.
• Removing tracing from a function takes more time with AOP than without. This is probably caused by the (first) usage of annotations.

• Selectively tracing parameters in a function, takes less time with AOP than without.

• Adding tracing to a function manually introduces significantly more errors than with AOP.

• Changing the signature of a function manually introduces significantly more errors than with AOP.

Besides this experiment, we are continuously working on the evaluation of the language and weaver by different means. At the moment of writing, software developers and engineers (approximately 70-80 participants) - who are currently using the language and weaver - are participating in a survey to evaluate the tool in the view of usability and other quality concerns. We consider the feedback both from the experiment and survey crucial in driving our design on the upcoming features of both the language and the weaver.

5.5 Conclusions

Currently, WeaveC is part of the standard build process of the ASML software. WeaveC is used in 42 components today; in these components 298 targets are generated based on 1007 woven source files. Software developers do not need to write tracing and timing code anymore. The aspect files that realize these crosscutting concerns are part of the external interface of a standard software component. The weaving of these aspects is enabled by make-files in the build process. Besides, developers can add annotations to their base code to customize the standard tracing functionality when needed.

Mirjam and WeaveC can fulfill the promises of the proof-of-concept weaver. The quality of the code ensured by Mirjam and WeaveC has positive impacts both on the maintenance effort and lead-time in the first line software development process. The increased quality of the code also improves lead-time and reduces errors in terms of the analysis of problems/machine performance, especially during integration and field problems. As a result, the forth line software development can also be done faster and better, which therefore reduces integration time and improves the response to field problems.
References


