Context-oriented Programming in ContextL

State of the Art

Pascal Costanza
Programming Technology Lab
Vrije Universiteit Brussel
B-1050 Brussels, Belgium
pascal.costanza@vub.ac.be

ABSTRACT
There is a wide range of scenarios where software systems have to be able to behave differently according to their context of use. In Context-oriented Programming (COP), programs can be partitioned into behavioral variations that can be freely activated and combined at runtime with well-defined scopes, such that the program behavior is affected depending on context. About four years ago, we have introduced our vision of Context-oriented Programming and have presented the programming language ContextL as an extension to the Common Lisp Object System (CLOS), as our first language extension that explicitly realizes this vision. Since then, ContextL has been picked up by various developers world-wide, is now in use in several software systems, and has been continuously improved to meet the demands of its users. For these reasons, ContextL can currently be regarded as the most mature realization of COP concepts. In this paper, we give an overview of the major ingredients of ContextL, describe the developments in ContextL of the last four years, and sketch some future work.

Categories and Subject Descriptors
D.1 [Software]: Programming Techniques—Object-oriented Programming; D.3.3 [Programming Languages]: Language Constructs and Features

Keywords
Context-oriented programming, behavioral variations, layer activation, dynamic scoping

1. CONTEXT-ORIENTED PROGRAMMING
There is a wide range of scenarios where software systems have to be able to behave differently according to their context of use [2, 3, 4, 6, 16, 23, 25, 28]. In Context-oriented Programming (COP), programs can be partitioned into behavioral variations that can be freely activated and combined at runtime with well-defined scopes. Such behavioral variations consist of partial definitions for typical program entities, like classes, methods, functions, procedures, and so on. The essential ingredients of Context-oriented Programming are as follows:

Context is any information that can be computationally accessed in a software system.

Behavioral variations describe the context-dependent behavior of a software system as increments on the underlying context-independent program definitions.

Layers group such behavioral variations as first-class entities that can be referenced in a program at runtime.

Layer activation is achieved by language constructs that ensure that such layers are added at runtime, such that the respective partial program definitions have an influence on the actual behavior of a program.

Scoping of layer activation and deactivation ensures that the behavioral variations are only effective for well-defined parts of a program, and for well-defined durations.

Context-oriented Programming focuses on programming constructs to enable grouping, referencing, and activation and deactivation of layers of behavioral variations. It should thus be seen as a complement to the (equally important) research on context acquisition and reasoning, where the focus is on sensing (low-level) context data and inferring (high-level) context information (for example [25]). Our definition of context is open and pragmatic, in that it treats any computationally accessible information as potential parameters for influencing the behavior of a program. In contrast, the widely cited definition for context of Dey at al. [17] states that “Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves.” That definition distinguishes between relevant and irrelevant information, which is important when modelling systems for context acquisition and context reasoning. It differs in this regard from our definition that is targeted at structured ways of how to affect program behavior depending on context, which should be addressed by general-purpose constructs and thus independent of the kinds of context information they may or may not depend on.
About four years ago, we have introduced our vision of Context-oriented Programming and have presented the programming language ContextL as an extension to the Common Lisp Object System (CLOS), as our first language extension that explicitly realizes this vision [8]. Since then we have implemented the underlying ideas as extensions for various other programming languages, like Smalltalk [19], Java [9], Ruby and Python, together with various use cases. During the course of these experiments, the essential ingredients of context, behavioral variations and scoped layer activation have proved to be stable corner stones.\(^1\)

On top of that, ContextL has been picked up by various developers worldwide since its public release as an open-source language extension, is now in use in several software systems in the “real world,” and has been continuously improved to meet the demands of its users. For these reasons, ContextL can currently be regarded as the most mature realization of COP concepts.\(^2\)

2. CONTEXTL: STATE OF THE ART

2.1 Essential Features

ContextL is an extension to the Common Lisp Object System (CLOS), and thus includes object-oriented features like classes with multiple inheritance, slots (fields) and methods with multiple dispatch. Like CLOS, ContextL is not based on message sending, but on generic functions [5].\(^3\)

For each of the core defining constructs in CLOS, there is a corresponding defining construct in ContextL. For defining classes, ContextL provides \texttt{define-layered-class} as an analogue to CLOS’s \texttt{defclass}. Likewise, ContextL provides \texttt{define-layered-function} and \texttt{define-layered-method} as analogues to CLOS’s \texttt{defgeneric} and \texttt{defmethod} for defining generic functions and methods.

Layers can be introduced in ContextL with \texttt{deflayer} and an associated name, as follows.

\begin{verbatim}
(deflayer layer-name)
\end{verbatim}

Such layers can then be further specified with partial class and method definitions, which can be explicitly associated with such layers, as follows.

\begin{verbatim}
(define-layered-class class-name
  [:in-layer layer-name]
  {\{superclass\}'}
  {\{slot-specification\}''}
  {\{class-option\}'})

(defmethod function-name [:in-layer layer-name]
  {\{method-qualifier\}' parameters method-body})
\end{verbatim}

There is always a root layer present that defines the context-independent behavior of a program, and class and method definitions can be associated with the root layer by either omitting the :in-layer specification, or by using \texttt{t} as the layer name.\(^4\)

By default, only the root layer is active at runtime, which means that only definitions associated with the root layer effect the behavior of a program. Other layers can be activated at runtime by way of \texttt{with-active-layers}, as follows.

\begin{verbatim}
(with-active-layers {\{layer-name\}'
  body)
\end{verbatim}

Such a layer activation ensures that all the named layers affect the program behavior for the dynamic extent of the enclosed program code (\texttt{body}). Layer activation is dynamically scoped: Both direct and indirect invocations of generic functions (methods) and slot accesses in the control flow of the layer activation are affected. Furthermore, layer activation is restricted to the current thread of execution in multi threaded Common Lisp implementations, to avoid race conditions and interferences between different contexts.

Furthermore, layers can be deactivated at runtime by way of \texttt{with-inactive-layers}, as follows.

\begin{verbatim}
(with-inactive-layers {\{layer-name\}'
  body)
\end{verbatim}

Such a layer deactivation ensures that none of the named layers affect the program behavior for the dynamic extent of the enclosed program code anymore. As with \texttt{with-active-layers}, layer deactivation is dynamically scoped and restricted to the current thread of execution.

Layer activations and deactivations can be nested arbitrarily in the control flow of the program. Both multiple activations and deactivations of the same layer are ignored, each layer is only active or inactive at most once. Furthermore, the order of layer activation determines method specificity when generic functions are invoked: Methods from more recently activated layers in the control flow of a program are executed before methods from less recently activated layers and the root layer.

The above constructs allow referencing layers at runtime, but only as second-class citizens. However, for each of the above constructs, there exist corresponding first-class constructs: The functions \texttt{ensure-layered-function}, \texttt{ensure-layered-method} and \texttt{ensure-layer} enable defining layered functions, layered methods and layers at runtime. Furthermore, \texttt{funcall-with-layer-context} and \texttt{apply-with-layer-context} enable activating and deactivating computed combinations of layers, and \texttt{adjoin-layer} and \texttt{remove-layer} enable computing such combinations of layers from first-class layer representations.

Finally, there are cases where it is useful that layers are activated or deactivated globally for all threads without dynamic scope, for example for interactive testing and development, or for program deployment (at system startup time). The functions \texttt{ensure-active-layer} and \texttt{ensure-inactive-layer} enable such global activation and deactivation of layers.

2.2 Efficient Layer Activation

Context-oriented Programming encourages continually changing program behavior. For example, we have illustrated a use case in [9] where side effects on graphical objects trigger updates of their representations on the screen. Context-oriented Programming can successfully be used in such a scenario to avoid invalid screen updates in situations where graphical objects are in intermediate states during type of all types, the standard output stream, and so on.

\(^1\)All this work has been performed in close collaboration with Robert Hirschfeld.

\(^2\)ContextL can be downloaded from http://common-lisp.net/project/closer/contextl.html

\(^3\)Note, however, that our experiments with similar extensions for other object-oriented programming languages show that COP is compatible with message sending as well.

\(^4\)The symbol \texttt{t} is traditionally used in Lisp dialects to denote “general” concepts, like the boolean truth value, the super-
composite changes. This use case is characterized by repeated activation and deactivation of a layer that is responsible for the necessary screen updates.

Such examples show that it must be possible to implement layer activation and deactivation efficiently in order to make Context-oriented Programming a sane option in a programmer’s toolbox. It is indeed not obvious that layer activation and deactivation can be efficient. However, we have found an implementation for ContextL on top of CLOS with competitive performance. The key ingredients are as follows.

- Layer activation and deactivation leads to a runtime composition of multiple, arbitrary layers. Such layer composition is realized internally in ContextL by reusing CLOS’s support for multiple inheritance. This can be achieved by representing composed layers as dynamically created classes that multiply inherit from other layers, on top of which they are activated and which are themselves represented as classes. Dynamic class creation is supported by way of the CLOS Metaobject Protocol [20], and although dynamic class creation is costly in terms of performance, such classes can be cached in a way such that each combination of a given sequence of layers needs to be created only once, and can be looked up very efficiently from then on.

- In ContextL, method selection and combination must depend on the current sequence of active layers. Since layer combinations are already represented as classes, we can internally represent the current sequence of active layers as a (prototypical) instance of such a combined class, and drive method selection and combination by implicitly passing such an instance as an argument to an appropriately specialized parameter that is added to each layered method internally. In this case, efficiency also comes from caches that allow for fast lookup of applicable methods, which is already provided by typical CLOS implementations [21].

- Layer activation and deactivation must modify the current sequence of active layers in such a way that the change is only visible in the current thread of execution, to ensure that they do not interfere with other threads in multithreaded Common Lisp implementations. To achieve this, the representation of the current sequence of active layers is bound to a dynamically scoped, thread-local variable, which are directly supported by Common Lisp and for which implementation alternatives with well-understood performance characteristics exist [1].

In [9], we discuss our overall implementation strategy in more detail, together with some benchmarks that indeed show a low overhead for layer activation and deactivation in terms of performance.

2.3 Reflective Layer Activation

Organizing a program into layers may lead to dependencies between layers: Layers may require each other’s presence, or may be mutually exclusive. For example, a typical use case for Context-oriented Programming is to separate the generation of different output formats (html, pdf, json, etc.) from the same document tree into several layers. Activating a layer for one output format should thus deactivate the layers for other output formats. In [10], we illustrated this idea with another example, where different tariffs for cell phone usage are separated into different layers, which must be mutually exclusive and, at the same time, all require the presence of another base tariff layer.

In principle, it is possible to make such dependencies explicit by simply activating and deactivating all the involved layers in all the places where layer activation or deactivation occurs. However, this turns out cumbersome. In [10], we introduce a reflective API for ContextL that can be used to make layer activations and deactivations automatically trigger other layer activations and deactivations behind the scenes. The essential idea is that each layer activation computes the resulting sequence of active layers by invoking the function \texttt{adjoin-layer-using-class}, and each layer deactivation computes the resulting sequence of active layers by invoking \texttt{remove-layer-using-class}. These two functions are exposed by the ContextL API, are themselves layered functions, and can thus themselves be specialized by user-provided layered methods. Such layered methods can inspect the layers to be activated or deactivated, inspect the current sequence of active layers, and ensure that other required layers are implicitly activated as well, other excluded layers are implicitly deactivated, or invalid layer activations or deactivations are rejected.

In [10], we discuss the details of ContextL’s reflective architecture, illustrate its use by giving examples, and show that it can be implemented without affecting the performance achieved by the implementation strategy discussed in [9] (see above).

2.4 Description of Layer Dependencies

In spite of the usefulness of reflective layer activation, implementing extensions on top of ContextL’s reflective API proves to be a complex task. In [12], we show a first step of how feature diagrams, or better the feature description language (FDL, [15]) as their textual counterpart, can be used to provide a high-level description of layer dependencies, such as requirement and exclusion dependencies. FDL can be added as an extension on top of ContextL’s reflective API without changing any of ContextL’s internal implementation details, and without affecting its essential efficiency characteristics.

We have also gained some first experience with using feature diagrams for the requirements elicitation and design phases of context-oriented systems. Context-Oriented Domain Analysis (CODA) [14] is a notation and methodology for describing both the context-independent behavior and the context-dependent behavioral variations of a system. This notation is based on feature diagrams, but extends it to include – potentially dynamically changing – conditions under which layers are required or excluded. We are currently working on mapping CODA diagrams to context-oriented programming languages such as ContextL.

2.5 Impact

Since its original release in 2005, ContextL has been picked up by several developers world-wide, and is now used in a number of software systems in the “real world.” ContextL is typically employed in web applications, where layers are
used to separate generation of different output formats from the same document tree, and to provide different views and modes to different kinds of users depending on their respective tasks. For a description of one exemplary case see http://p-cos.blogspot.com/2007_11_01_archive.html.

Furthermore, Context-oriented Programming has started to impact other researchers who have picked up the term and the concepts as starting points for their own work [7, 13, 18, 26]. A detailed discussion of such related work is outside the scope of the overview given here.

3. FUTURE WORK AND CONCLUSIONS

There are several ways in which ContextL can be extended to broaden its scope of applicability:

- ContextL’s reflective API can be extended to support first-class dynamic environments. The idea is to support capturing the current sequence of active layers as a first-class entity such that it can be re-instantiated later for a different context. For example, this would allow passing context from one thread to another and storing the current context of a suspended computation to be later resumed, for example in continuation-based web applications [24]. Since frameworks for such continuation-based web applications are typically based on partial continuations, we actually need support for delimited dynamic bindings, as in [22]. Since there are different continuation frameworks available for Common Lisp, we have to design the reflective API for first-class dynamic environments in such a way that they are compatible with different kinds of partial continuations.

- Currently, layers in ContextL contain partial class and method definitions, where the method definitions are typically specialized on classes. Although CLOS supports method specialization on single objects as well, somewhat similar to what is provided in prototype-based object systems, the actual use of layer activation or deactivation for single objects is not straightforward, but would sometimes be useful. After some experimentation with different ideas, we have come up with a generalization of generic function dispatch called filtered dispatch which allows for filtering arguments before method selection and combination, while passing the original unfiltered arguments to the thus selected methods for execution [11]. This results in a very powerful and expressive dispatch mechanism, which has a potential to serve as a basis for context-dependent behavior of single objects.

- Context-dependent behavior is also important in distributed settings, especially in the case of ad-hoc mobile networks between mobile devices. Some first experiments in using COP concepts in such a setting are very promising. Interesting research topics are how to distribute layer activation across several nodes of such a network, and how to deal with potentially different context parameters from different interacting nodes. We have described first steps in that regard in the Context-Dependent Role Model [27], and are currently working on integrating a full-fledged distributed architecture with ContextL.

Other important areas of research need to address questions such as:

- What are good software architectures that can take advantage of Context-oriented Programming? There is now a considerable amount of practical experience with languages like ContextL, and some design principles have started to emerge, but they have not been codified yet, for example in the form of design patterns and architectural patterns, or in the form of libraries and frameworks that take advantage of Context-oriented Programming.

- What are good methodologies for eliciting requirements for, and designing context-oriented software systems? We have made some first steps with CODA [14] (see above), and current research in the field of Software Product Lines and especially Dynamic Software Product Lines are very promising starting points.

- How can context-aware systems be tested and verified? This is an inherently hard issue because of the combinatorial explosion of possibilities induced by the presence of multiple context parameters, which is reflected in Context-oriented Programming by the presence of multiple layers that can be arbitrarily combined with each other in the general case.

At this stage in time, we can rightfully claim that Context-oriented Programming has reached a certain level of maturity, has fostered some interesting research results, and has started to gain traction in industry. ContextL, as based on the Common Lisp Object System, has proved to be a viable tool both for research and for industrial use, and will therefore remain one of our main vehicles for exploring the field in more depth.

Acknowledgments

The central concepts of Context-oriented Programming have been discovered and defined in close collaboration with Robert Hirschfeld over the years.


4. REFERENCES


