

Unclassified, Draft

A Formal Model of OSGi R4 Modularity (v0.83)

Glyn Normington

September 19, 2007

The aim is to model a subset of the features of the modularity layer of OSGi Release 4. This is a (stalled) work in progress.

1 Introduction

This specification only models a subset of the modularity function of OSGi R4.
It does not model:

- ‘uses’
- optional resolution
- dynamic imports
- export filters
- require-module

2 UML Overview

Figure 1 shows the main artefacts of OSGi modularity, some of which will be covered by the formal model.

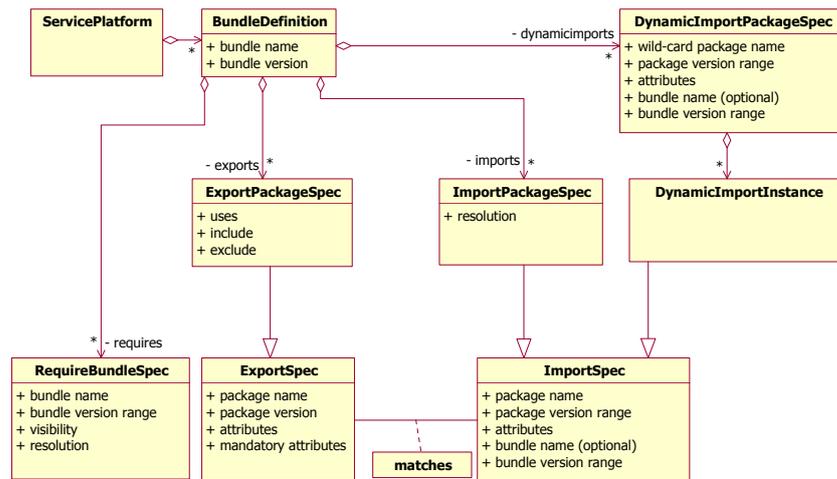


Figure 1: UML Overview

There are three new concepts in the diagram which are purely to explain the other “classes” which relate directly to RFC 79 manifest headers.

ExportSpec is a piece of the “export signature” of a module which is a generalisation of `Export-Package`. What the diagram doesn’t show is that `Require-Bundle` with `visibility:=reexport` also produces these abstract exports.

ImportSpec is a kind of generalisation of `Import-Package` and `DynamicImport-Package`. `AbstractImport` really does generalise `Import-Package`. But `DynamicImport-Package` can, because of its wild-carded package name, be thought of as a kind of macro or template which generates (at class load time) “dynamic import instances”, described next.

DynamicImportInstance is effectively an import which is derived, at class load time, from a dynamic import. `AbstractImport` generalises `DynamicImportInstance`.

The crux of the diagram is that there is a “matches” relationship between `AbstractImport` and `AbstractExport`. This relationship describes potential package wirings - RFC 79 has to cover the cases when an import matches more than one export.

3 Basic Types

Some basic types need defining.

Module Names

Modules are identified within a module system by module name and module version number.

$$[MName, MVer]$$

Module version, consisting of the name and version number of a module, needs an abbreviation.

$$MV == MName \times MVer$$

Packages

We need the notion of packages (or, more precisely, package names) and package version numbers.

$$[Package, PVer]$$

Package version, consisting of the name and version number of a package, needs an abbreviation.

$$PV == Package \times PVer$$

We need the notion of (fully qualified) class names and classes.

$$[ClassName, Class]$$

Each class name belongs a unique package.

$$| \text{ package} : ClassName \rightarrow Package$$

Certain classes belong to packages beginning with “java.” (and “org.omg.” etc.).

$$| \text{ javaClasses} : P \text{ } ClassName$$

Class Spaces and Consistency

A class space is a collection of classes indexed by class name.

$$ClassSpace == ClassName \text{ } \text{ } Class$$

We say that two class spaces s and t are *consistent* if and only if

$$s \cup t \in ClassSpace$$

, i.e. s and t agree on the class names they have in common.

Attributes

Arbitrary attribute names and values will be used to provide a flexible import/export matching mechanism.

$$[AName, AValue]$$

We will use mappings of attribute names to values.

$$AMap == AName \rightarrow AValue$$

Module names are sometimes used as attribute names.

$$| MNameAttr : AName$$

4 Simple Class Loaders

Before we continue to describe OSGi modularity, we model a simple class loader which delegates to a parent class loader. This will help to familiarise the reader with the notation and how class loading is modelled.

A simple class loader has a local collection of class definitions, typically implemented as a sequence of jar files on a file system, but modelled here as a class space. It also has a class space of classes which have been defined (i.e. loaded locally) by the class loader and which are derived from the loader's local collection of class definitions. Finally it has a class space of all classes loaded by the class loader. We defer tying this to the the parent class loader.

<i>SimpleLoader</i> <i>definitions</i> : <i>ClassSpace</i> <i>defined</i> : <i>ClassSpace</i> <i>loaded</i> : <i>ClassSpace</i>
$defined = definitions \cap loaded$

The defined classes are those which have definitions and which are also loaded. Note that this only makes sense because the child and parent definitions turn out to be disjoint. If a class is loaded from identical class files by two distinct class loaders, the resultant class objects are distinct.

A simple loader initially has no loaded classes.

<i>SimpleLoaderInit</i> <i>SimpleLoader</i>
$loaded = \emptyset$

We now model a successful class load.

<i>SuccessfulLoad</i> Δ <i>SimpleLoader</i> <i>cn?</i> : <i>ClassName</i>
$definitions' = definitions$ $defined \subseteq defined'$ $loaded \subseteq loaded'$ $cn? \in \text{dom } loaded'$

The definitions are preserved. The collections of defined and loaded classes may not decrease. The specified class is loaded. Note that more than one class may be loaded, which may be necessary if the specified class refers to other classes which have not yet been loaded.

A *bootstrap* class loader is at the top of the parent-child hierarchy and therefore only loads classes it also defines. We refer to such a class loader as a *top loader*.

<i>TopLoader</i>
<i>SimpleLoader</i>
$defined = loaded$

We shall see later that a simple loader together with its parentage up to and including the bootstrap loader may be modelled as a top loader.

With this in mind, we can express the relationship between a child class loader and its parentage.

<i>ParentalConstraint</i>
$child : SimpleLoader$
$parentage : TopLoader$
$child.loaded \setminus child.defined \subseteq parentage.loaded$
$(dom\ child.defined) \cap (dom\ parentage.definitions) = \emptyset$
$child.definitions \cap parentage.definitions = \emptyset$

Classes which the child has loaded but not defined must have been loaded by the parentage. The child has not defined a class for which the parentage has a definition. The child and parentage definitions are disjoint which models the fact that distinct class loaders produce distinct class instances when they load from the same class bytecode.

The first two properties are ensured by the child class loader first of all delegating each class load to its parent and then only attempting to define the class if its parent failed to load the class.

It follows that $child.loaded$ and $parentage.loaded$ are consistent class spaces. The proof is omitted because of lack of space.

We now show how a child together with a top loader satisfying the above constraint can be combined to produce a new top loader.

<i>Combine</i>
<i>ParentalConstraint</i>
$loader : SimpleLoader$
$loader.definitions = child.definitions \oplus parentage.definitions$
$loader.defined = child.defined \cup parentage.defined$
$loader.loaded = child.loaded \cup parentage.loaded$

The proof that the result is really a top loader is omitted for lack of space.

$Combine \vdash loader \in TopLoader$

The reverse process of factoring a child class loader from a top loader is also sometimes possible, but is not explored further here.

5 Delegating Class Loaders

As a further step towards OSGi R4 modularity, we extend the notion of a simple loader to model imports and exports.

An importer is a simple loader which defines the class names it is willing to import from elsewhere. In OSGi R4 as we shall see later, imports are defined in terms of packages, but we will overlook that level of detail for the moment. For simplicity, we also ignore the parent class loader.

<i>Importer</i> <i>SimpleLoader</i> <i>imports</i> : P <i>ClassName</i> <i>imported</i> : <i>ClassSpace</i>
<i>imported</i> = <i>loaded</i> \ <i>defined</i> $\text{dom } \textit{imported} \subseteq \textit{imports}$

Classes which are loaded, but not defined, are imported. Imported classes must be named as imports.

An exporter is a simple loader which defines the class names it is willing to export to an importer.

<i>Exporter</i> <i>SimpleLoader</i> <i>exports</i> : P <i>ClassName</i> <i>exported</i> : <i>ClassSpace</i>
<i>exported</i> \subseteq <i>loaded</i> $\text{dom } \textit{exported} \subseteq \textit{exports}$

Exported classes must be loaded and named as exports.

We now define a constraint between a ‘matching’ importer and exporter.

<i>DelegationConsistency</i> <i>i</i> : <i>Importer</i> <i>e</i> : <i>Exporter</i>
<i>i.definitions</i> \cap <i>e.definitions</i> = \emptyset let <i>shared</i> == <i>i.imports</i> \cap <i>e.exports</i> • $\textit{shared} \subseteq \textit{i.imported} \subseteq \textit{shared} \subseteq \textit{e.exported}$

The importer and exporter definitions are disjoint which models the fact that distinct class loaders produce distinct class instances when they load from the same class bytecode. If shared classes denote those which may be both imported by the importer and exported by the exporter, then the shared classes which have been imported have also been exported.

It follows that $\textit{shared} \subseteq \textit{i.loaded}$ and $\textit{shared} \subseteq \textit{e.loaded}$ are consistent class spaces.

One of the main problems solved by OSGi R4 is how to match imports to exports in order to maintain proper consistency of class spaces and avoid certain kinds of type mismatches resulting in class loading failures, class cast exceptions, etc.

The solution starts with a ‘module definition’.

6 Module Definition

A module system contains a collection of modules which share packages in various ways.

Before we model the module system, we focus first on the definition of an individual module and secondly, in the next section, on the runtime state of an individual module. A module's definition identifies the module by module name and module version. It also contains class definitions which are contained in the module and may be loaded by the module.

<i>ModuleId</i> <i>mname</i> : <i>MName</i> <i>mver</i> : <i>MVer</i> <i>mv</i> : <i>MV</i>
<i>mv</i> = (<i>mname</i> , <i>mver</i>)

<i>ModuleDef₀</i> <i>ModuleId</i> <i>definitions</i> : <i>ClassSpace</i>
<i>mv</i> = (<i>mname</i> , <i>mver</i>)

A complete list of imported and exported classes would be hard to maintain. Also a given class module does not typically want an arbitrary mixture of classes from various sources. The OSGi design point is to make the Java package the minimum unit of resolution.

Each package which a module exports has an export specification which specifies the module name, module version, package version, and a set of matching attributes some of which must be supplied by an importer which wishes to use the exported package. The export specification may also require a matching importer to specify module name or module version (or both).

<i>ExportSpec</i> <i>ModuleId</i> <i>pver</i> : <i>PVer</i> <i>attr</i> : <i>AMap</i> <i>mandatory</i> : P <i>AName</i>
<i>MNameAttr</i> \notin dom <i>attr</i> <i>mandatory</i> \subseteq (dom <i>attr</i> \cup { <i>MNameAttr</i> })

Each package which a module imports has an import specification.

A module name may be specified (using a singleton set) or not (using an empty set). Module versions may only be constrained if a module name has been specified. A range of acceptable module versions or package versions is represented as a set containing all the elements of the range.

$ \begin{array}{l} \textit{ImportSpec} \\ \hline mname : P \textit{MName} \\ mver : P \textit{MVer} \\ pver : P \textit{PVer} \\ attr : AName \rightarrow AValue \\ \hline \#mname \leq 1 \\ mname = \emptyset \Rightarrow mver = \textit{MVer} \\ \textit{MNameAttr} \notin \text{dom } attr \end{array} $

An export specification matches an import specification if and only if:

- any module name, module versions, and package versions in the import specification are satisfied by exported module name, module version, and package version,
- any arbitrary attributes specified in the import specification match the values specified in the export specification,
- any arbitrary attributes specified as mandatory in the export specification are specified in the import specification,
- if the export specification mandated the module name, the import specification specifies the module name.

$ \begin{array}{l} matches : \textit{ImportSpec} \leftrightarrow \textit{ExportSpec} \\ \hline \forall is : \textit{ImportSpec}; es : \textit{ExportSpec} \bullet \\ is \mapsto es \in matches \Leftrightarrow \\ \quad is.mname = \emptyset \vee is.mname = \{es.mname\} \wedge \\ \quad es.mver \in is.mver \wedge \\ \quad es.pver \in is.pver \wedge \\ \quad is.attr \subseteq es.attr \wedge \\ \quad es.mandatory \subseteq (\text{dom } is.attr \cup \{\textit{MNameAttr}\}) \wedge \\ \quad \textit{MNameAttr} \in es.mandatory \Rightarrow is.mname \neq \emptyset \end{array} $
--

A module's definition identifies the packages the module imports and exports by providing corresponding import and export specifications. A module may provide at most one import specification for a given package name but it may have multiple export specifications for a given package name.

A module may not name itself in an import specification.

However a module can import and export the same package, although as we shall see later, such a package is defined by one and only one module. If a module imports and exports the same package, either the import or the export is honoured when the module is resolved and the other statement is disregarded. We add abbreviations for the sets of classes which may be exported and imported.

ModuleDef

ModuleDef₀

exports : *Package* \leftrightarrow *ExportSpec*

imports : *Package* \rightarrow *ImportSpec*

classExports, *classImports* : P *ClassName*

$\forall es : \text{ran } exports \bullet es.mv = mv$

$\forall is : \text{ran } imports \bullet mv \notin is.mname \times is.mver$

classExports = *package*[~](dom *exports*)

classImports = *package*[~](dom *imports*)

7 Module

Based on its definition, a module loads classes either from its contents or by importing from another module¹. These loaded classes are modelled using various class spaces in addition to the module definition:

defined classes loaded by the module²,

imported classes imported from other modules.

Note that *defined* and *imported*, together with the module's definition, determine the other class spaces:

loaded all classes available to a module,

exported classes exported to other modules,

private classes loaded by the module which are not exported.

<p><i>Module</i></p> <hr/> <p><i>ModuleDef</i></p> <p><i>defined</i> : <i>ClassSpace</i></p> <p><i>imported</i> : <i>ClassSpace</i></p> <p><i>loaded</i> : <i>ClassSpace</i></p> <p><i>exported</i> : <i>ClassSpace</i></p> <p><i>private</i> : <i>ClassSpace</i></p> <hr/> <p><i>defined</i> \subseteq <i>definitions</i></p> <p><i>defined</i> = <i>classImports</i> $\dot{\cup}$ <i>loaded</i></p> <p><i>imported</i> = <i>classImports</i> $\dot{\cap}$ <i>loaded</i></p> <p><i>exported</i> = <i>classExports</i> $\dot{\cap}$ <i>loaded</i></p> <p><i>private</i> = <i>classExports</i> $\dot{\cup}$ <i>loaded</i></p>

Some interesting properties follow.

$$\begin{aligned}
 \text{Module} \vdash & \\
 & \langle \text{defined}, \text{imported} \rangle \text{ partition } \text{loaded} \wedge \\
 & \langle \text{private}, \text{exported} \rangle \text{ partition } \text{loaded} \wedge \\
 & \text{package}(\text{dom } \text{defined}) \cap \text{dom } \text{imports} = \emptyset \wedge \\
 & \text{defined} = \text{definitions} \cap (\text{loaded} \setminus \text{imported})
 \end{aligned}$$

The defined and imported class spaces partition the loaded class space. The private and exported class spaces also partition the loaded class space. No classes are loaded locally which belong to imported packages. Locally loaded classes are precisely those which have local definitions and which have been loaded but not imported.

The proofs of these properties is left as an exercise for the reader.

¹For the purposes of this model, we ignore the parent class loader

²Strictly speaking, *defined* models classes for which the module's class loader is the *defining* loader.

8 Module System

To give an indication of how the pieces of the specification fit together, we model a service platform of installed modules (m) some of which are resolved and with a wiring function ($wire$) wiring together resolved modules for packages.

<i>ModuleSystemBase</i>
$m : MV \text{ } \text{æ} \text{ } Module$ $resolved : P \text{ } MV$ $wire : MV \times Package \rightarrow MV$
$ran \text{ } wire \subseteq resolved$ $resolved \subseteq dom \text{ } m$ $first \text{ } (dom \text{ } wire) \subseteq resolved$ $(\forall mv : resolved \mid (m \text{ } mv).imports \neq \emptyset \bullet mv \in first \text{ } (dom \text{ } wire))$

Note that the package wiring for a given package p is described by the function $\lambda mv : MV \bullet wire(mv, p)$.

Two constraints must be satisfied. The first is that imports must be matched by exports.

<i>ImportExportMatching</i>
<i>ModuleSystemBase</i>
$(\forall m1 : MV; p : Package \mid (m1, p) \in dom \text{ } wire \bullet$ $p \in dom(m \text{ } m1).imports \wedge$ $p \in dom((m \circ wire)(m1, p)).exports \wedge$ $(let \text{ } is == (m \text{ } m1).imports \text{ } p \bullet$ $(\exists es : (((m \circ wire)(m1, p)).exports) \{ \{p\} \} \bullet$ $(is, es) \in matches)))$

The second is that matched importers and exporters are consistent.

<i>WiringConsistency</i>
<i>ModuleSystemBase</i>
$\forall i, e : MV; p : Package \mid (i, p) \mapsto e \in wire \bullet$ $let \text{ } mi == m \text{ } i; me == m \text{ } e; c == package^{\sim} \{ \{p\} \} \bullet$ $c \text{ } C \text{ } mi.imported \subseteq c \text{ } C \text{ } me.exported$

Then a module system must satisfy both constraints.

$$ModuleSystem \hat{=} ImportExportMatching \wedge WiringConsistency$$

<i>ResolveOk</i>
$\Delta ModuleSystem$
$m? : MV$ $mdef? : ModuleDef$
$m? \notin dom \text{ } m$ $m? \in resolved'$

9 Class Search Order

Although we have avoided modelling the parent class loader and Require-Bundle so far, it is essential to see how these features combine into the final class search order. We add the parent class loader in, modelled as a top loader, and the contribution of all required modules modelled as a class space. We also model the search order explicitly as a class space.

<i>CompleteModule</i>
<i>parent</i> : <i>TopLoader</i> <i>required</i> : <i>ClassSpace</i> <i>search</i> : <i>ClassSpace</i> <i>ModuleDef</i> <i>defined</i> : <i>ClassSpace</i> <i>imported</i> : <i>ClassSpace</i> <i>loaded</i> : <i>ClassSpace</i> <i>exported</i> : <i>ClassSpace</i> <i>private</i> : <i>ClassSpace</i>
$(classImports \cup classExports) \cap javaClasses = \emptyset$ $classImports \cap \text{dom } required = \emptyset$ $loaded = (defined \cup required \cup imported)$ $\quad \oplus (javaClasses \text{ C } parent.loaded)$ $search = ((classImports \notin (definitions \oplus required)) \cup imported)$ $\quad \oplus (javaClasses \text{ C } parent.definitions)$ $defined \subseteq definitions$ $defined = ((classImports \cup javaClasses) \notin loaded) \setminus required$ $imported = classImports \text{ C } loaded$ $exported = classExports \text{ C } loaded$ $private = classExports \notin loaded$

Neither imports nor exports may specify packages beginning with “java.”. Required classes do not include any classes which are specified as imports.

Loaded classes consist of those defined by the module, those required from other modules, those imported from other modules, but all overridden with the parent class loader’s packages which begin with “java.”.

The search order reflects the parent class loader being searched for packages beginning with “java.”, and then imported packages, and then required and defined classes which do not belong to imported packages.

The defined classes are those which are loaded but not inherited from the parent, imported, or required. Imported classes are those which are loaded and which belong to imported packages. Exported classes are those which are loaded and which belong to exported packages. Private classes are those which are loaded but which do not belong to exported packages.

10 Z Notation

Numbers:

\mathbb{N} Natural numbers $\{0, 1, \dots\}$

Propositional logic and the schema calculus:

$\dots \wedge \dots$	And	$\langle \dots \rangle$	Free type injection
$\dots \vee \dots$	Or	$[\dots]$	Given sets
$\dots \Rightarrow \dots$	Implies	$\prime, ?, !, 0 \dots 9$	Schema decorations
$\forall \dots \mid \dots \bullet \dots$	For all	$\dots \vdash \dots$	theorem
$\exists \dots \mid \dots \bullet \dots$	There exists	$\theta \dots$	Binding formation
$\dots \setminus \dots$	Hiding	$\lambda \dots$	Function definition
$\dots \cong \dots$	Schema definition	$\mu \dots$	Mu-expression
$\dots == \dots$	Abbreviation	$\Delta \dots$	State change
$\dots ::= \dots \mid \dots$	Free type definition	$\Xi \dots$	Invariant state change

Sets and sequences:

$\{\dots\}$	Set	$\dots \setminus \dots$	Set difference
$\{\dots \mid \dots \bullet \dots\}$	Set comprehension	$\bigcup \dots$	Distributed union
$\mathbb{P} \dots$	Set of subsets of	$\# \dots$	Cardinality
\emptyset	Empty set	$\dots \subseteq \dots$	Subset
$\dots \times \dots$	Cartesian product	$\dots \subset \dots$	Proper subset
$\dots \in \dots$	Set membership	$\dots \text{ partition } \dots$	Set partition
$\dots \notin \dots$	Set non-membership	seq	Sequences
$\dots \cup \dots$	Union	$\langle \dots \rangle$	Sequence
$\dots \cap \dots$	Intersection	disjoint ...	Disjoint sequence of sets

Functions and relations:

$\dots \leftrightarrow \dots$	Relation	$\dots \mapsto \dots$	maplet
$\dots \rightarrow \dots$	Partial function	$\dots \sim$	Relational inverse
$\dots \twoheadrightarrow \dots$	Total function	\dots^*	Reflexive-transitive closure
$\dots \ni \dots$	Partial injection	$\dots (\dots)$	Relational image
$\dots \ni \dots$	Injection	$\dots \oplus \dots$	Functional overriding
dom...	Domain	$\dots \mathbb{C} \dots$	Domain restriction
ran...	Range	$\dots \mathbb{C} \dots$	Domain subtraction

Axiomatic descriptions:

<i>Declarations</i>
<i>Predicates</i>

Schema definitions:

<i>SchemaName</i>
<i>Declaration</i>
<i>Predicates</i>