

# 1 **Position Paper: Modularity,**

# 2 **Namespaces and Versioning**

3 **Author:** (see Change History section)

4 **Date:** 2-26-2003

5 **Filename:** draft-burcham-modnamver-07.doc

6

7	1	Change History .....	3
8	2	Summary .....	3
9	3	Problem Description .....	3
10	4	Assumptions.....	4
11	4.1	Problem Size .....	4
12	4.2	Optimal Component Size.....	4
13	5	Options: XML Namespace Identification.....	5
14	5.1	Option 1: Namespace Name = Namespace Location .....	5
15	5.2	Option 2: Namespace Name is OASIS URN namespace .....	5
16	6	Recommendation: Namespace Identification .....	5
17	7	Recommendation: Schema Location .....	5
18	8	Options: Namespace Structure.....	6
19	8.1	Option 1: One Big Namespace .....	6
20	8.2	Option 2: One Namespace Per Type.....	6
21	8.3	Option 3: Core Plus “Functional” Namespaces .....	6
22	8.4	Option 4: Core Plus “Functional” Namespaces Plus Internal Structure as	
23		Needed .....	7
24	9	Recommendation: Namespace Structure .....	7
25	9.1	Into What Namespace Do Extensions Go.....	7
26	10	Options: Module Structure.....	8
27	10.1	Option 1: One Module Per Namespace .....	8
28	10.2	Option 2: One Module Per Object Class.....	8

29	10.3	Option 3: Modules based on Human Judgment of Related Functionality of	
30		Type Definitions .....	8
31	10.4	Option 4: One Module per Type Definition .....	9
32	11	Recommendation: Module Structure .....	9
33	11.1	Message Types.....	11
34	11.2	Number of Message Types .....	11
35	12	Options: Versioning.....	12
36	12.1	Option XF-1: Change the (internal) schema “version” attribute .....	12
37	12.2	Option XF-2: Create a “schemaVersion” attribute on the root element .....	12
38	12.2.1	Usage A: Conformance enforced by validator.....	12
39	12.2.2	Usage B: Conformance enforced by an extra processing pass .....	12
40	12.3	Option XF-3: Change the schema’s target namespace .....	12
41	12.4	Option XF-4: Change the name/location of the schema.....	12
42	12.5	Option 5: Schema Version as Context Classifier.....	12
43	13	Recommendations: Versioning.....	12
44	14	Definitions.....	16
45	15	References.....	16
46			

# 47 1 Change History

Revision	Editor	Description
0.7	Bill Burcham	Fleshed-out Recommendations:Versioning. Finalized namespace URI's throughout.
0.6	Dave Carlson	Changes made to section "Options: Module Structure" to enumerate several options for module definition.
0.1-0.5	Bill Burcham	Baseline

48

## 49 2 Summary

50 There are many possible mappings of XML schema constructs to namespaces and to  
51 operating system files. This paper explores some of those alternatives and sets forth  
52 some rules governing that mapping in UBL.

## 53 3 Problem Description

54 Namespaces are a syntactic convenience supporting the association of a "context" with  
55 either a lexical scope (default namespace), or a shorthand identifier (namespace  
56 qualifier). This context, applied either implicitly (in a lexical scope) or explicitly (via  
57 qualified names) supports compression of what would otherwise be long identifiers. In  
58 the absence of namespaces, identifier names are all long.

59 It is common for an instance document to carry namespace declarations, so that it might  
60 be validated. Processing logic (such as a stylesheet) typically carries namespace  
61 declarations pertaining to the language in which it is specified in (XSLT) as well as the  
62 namespaces on which it *operates*. The latter must match namespaces in the instance  
63 document under translation in order for useful work to be carried out.

64 In practice, namespaces are often given names denoting a hierarchy. XML processing  
65 tools may or may not use this hierarchy information. This sort of hierarchical naming  
66 though can be useful for the human reader.

67 As with other significant software artifacts, schemas can become large. In addition to the  
68 logical taming of complexity that namespaces provide, we might like to also divide the  
69 physical realization of that schema into multiple operating system files.

70 Schemas change over time. UBL will be no exception. What sort of version information  
71 (if any) will a schema carry? How shall that information be carried so as to conveniently  
72 support the needs of users operating on document instances with XML processing tools.

73 This position paper will address these three topics related to namespaces:

- 74 1. **Namespace Structure:** What shall be the mapping between namespaces and  
75 XML Schema constructs (e.g. type definitions)?
- 76 2. **Module Structure:** What shall be the mapping between namespaces and XML  
77 Schema constructs and operating system files?

78 3. **Versioning**: What support for versioning of schema shall be provided?  
79 In subsequent sections, we'll examine each topic in turn, presenting first the options, then  
80 a recommendation.

## 81 **4 Assumptions**

82 Much of this discussion will be based on the expected complexity of the UBL  
83 vocabulary. We structure systems into components in order to manage complexity.

### 84 **4.1 Problem Size**

85 How big will UBL be? How interconnected?

86 One source for complexity estimation is xCBL. **TBD: how many type definitions,**  
87 **element declarations, "instance roots" in xCBL?**

88 Another source for estimation is X12 that according to [NDR-MSG-88] has:

89 a bit over 1,000 data elements (...) a smaller number of segments, and  
90 300 or so transaction sets

91 Also from [NDR-MSG-88] we have EDIFACT:

- 92 ▪ There are just under 650 data elements which are
- 93 ▪ used in approx 200 composite structures (sort of equivalent to low level
- 94 Aggregate Core Components (ACCs)).
- 95 ▪ These elements and composites are reused within just over 150 segment
- 96 structures (sort of equivalent to higher level ACCs).
- 97 ▪ Combinations of all the above make up just under 200 messages (doc
- 98 types).

99 So an estimate of 1000 types and 250 message types seems reasonable for UBL.

### 100 **4.2 Optimal Component Size**

101 We don't want to define 1000 types all in one XML namespace, nor would we want to  
102 define them all in one file. Such an approach would lack structure necessary for  
103 understanding both by maintainer and users. Additionally, performance would be far  
104 from optimal for instance documents that needed only a subset of the UBL types.

105 For these reasons we presume that we need to structure and divide UBL into a hierarchy  
106 of components. We will strive to balance coupling and cohesion between the  
107 components in order to:

- 108 ▪ Manage the complexity of each component while not creating too many  
109 components<sup>1</sup>

---

<sup>1</sup> The "seven plus or minus two" rule [SEVEN-TWO] is a good, general rule of thumb. It's especially useful when you don't have any other rule. It says that if you want people to be able to keep a set of concepts in mind, then you are limited to about seven concepts.

110       ▪ Provide for useful subsetting of components  
111 We envision that many useful instance documents (messages) will be possible that  
112 require only a fraction of the overall UBL schema. In those cases it should be possible to  
113 avoid processing of the unneeded parts.

## 114 **5 Options: XML Namespace Identification**

115 This section presents some options for the form that UBL namespace names might take.

### 116 **5.1 Option 1: Namespace Name = Namespace Location**

117 There is certainly precedent for this approach. See for example the ebXML Message  
118 Service schema [http://www.oasis-open.org/committees/ebxml-msg/schema/msg-header-2\\_0.xsd](http://www.oasis-open.org/committees/ebxml-msg/schema/msg-header-2_0.xsd).

### 119 **5.2 Option 2: Namespace Name is OASIS URN 120 namespace**

121 This option exemplifies the current best practice within OASIS. See RFC 3121 [OASIS-  
122 URN-NS] for details. See Namespaces in XML for background [NAMESPACE].

123 Under this option, the namespace names for UBL namespaces would have the following  
124 form while the schemas are at draft status:

125 `urn:oasis:names:tc:ubl:schema{:subtype}?:{document-id}`

126 When they move to specification status the form will change to:

127 `urn:oasis:names:specification:ubl:schema{:subtype}?:{document-id}`

128 Where the form of {document-id} is **TBD** but should match the schema module name  
129 (see section 7, Recommendation: Schema Location).

## 130 **6 Recommendation: Namespace 131 Identification**

132 We pick Option 2: *Namespace Name is OASIS URN namespace*.

133 Will document-id include versioning information or will versioning be handled outside  
134 this identifier? See section 13, Recommendations: Versioning.

## 135 **7 Recommendation: Schema Location**

136 A question related to Namespace identification is schemaLocation. Schema location  
137 includes the complete URI which is used to identify schema modules.

---

Implications for XML for example might be: a type would define no more than seven (or so) elements, a namespace would define no more than about seven types, etc.

138 In the fashion of other OASIS specifications, UBL schema modules will be located under  
139 the UBL committee directory:

140 <http://www.oasis-open.org/committees/ubl/schema/<schema-mod-name>.xsd>

141

142 **TBD** does this recommendation need more justification?

143 Where <schema-mod-name> is the name of the schema module file. The form of that  
144 name is **TBD**.

145 There are two issues here. One is: how do we tell *users* to reach our schemas and two:  
146 what do we use internally – URN’s or URL’s. One is where/how do we publish our  
147 schemas.

## 148 **8 Options: Namespace Structure**

149 In this section we’ll explore some mappings between XML Schema structures and  
150 namespaces.

### 151 **8.1 Option 1: One Big Namespace**

152 We could have one big namespace for UBL. On the plus side, it would be fairly easy to  
153 remember. The downside is that we would forfeit the opportunity to use hierarchical  
154 namespaces to communicate the structure of the vocabulary.

### 155 **8.2 Option 2: One Namespace Per Type**

156 This approach represents the other end of the spectrum. If you’ve got a namespace per  
157 type then why not just use the type name. The namespace fails to be shorthand for  
158 anything. It fails to be memorable, or to group related types together.

### 159 **8.3 Option 3: Core Plus “Functional” Namespaces**

160 This option represents a space between 8.1 and 8.2. There would be namespaces for  
161 “core” types and there would be namespaces for each of the functional areas e.g. Order,  
162 Invoice.

Purpose	Namespace name
Common Leaf Types	urn:oasis:names:tc:ubl:schema:CommonLeafTypes:major-version:minor-version
Common Aggregate Types	urn:oasis:names:tc:ubl:schema:CommonAggregateTypes:major-version:minor-version
Order Domain	urn:oasis:names:tc:ubl:schema:Order:major-version:minor-version
Invoice Domain	urn:oasis:names:tc:ubl:schema:Invoice:major-version:minor-version
<b>TBD</b>	<b>TBD</b>

163

164 This represents a top-level decomposition of the vocabulary into multiple vertical  
165 (functional) slices: Order, Invoice; and two (horizontal) slices – the so-called core,  
166 CommonLeafTypes and CommonAggregateTypes.

167 The downside of this approach is that with seven or so functional namespaces, they are  
168 going to get awfully “crowded” (on the order of one hundred types per namespace).

## 169 **8.4 Option 4: Core Plus “Functional” Namespaces Plus** 170 **Internal Structure as Needed**

171 A refinement on 8.3, this option frees each of the functional and core namespaces to have  
172 their own hierarchy as necessary in order to further manage complexity.

173 [Add explanation from the minutes here.](#)

## 174 **9 Recommendation: Namespace Structure**

	Pro	Con
Option 1: one big namespace	Easy to remember namespace	When anything in UBL changes, all processing code must be changed (at a minimum to use new namespace name)
Option 2: namespace per type	Total compartmentalization	Why use namespaces at all? With this option the namespace ceases to provide useful contextualization.
Option 3: core plus “functional” namespaces	Allows parts of UBL to change independently. When a functional area changes, processing code depending on core needn’t change.	Doesn’t allow for intermediate structure. What if the functional namespaces may require further subdivision?
Option 4: core plus “functional” namespaces plus internal structure as needed	(same as Option 3)	By allowing intermediate namespaces, they will certainly flourish. Design rules must be developed to avoid regressing toward Option 2 over time.

175

176 [Option 3 is recommended. We reserve the right to revisit this decision when we are](#)  
177 [further along in the process of defining types. If we find that we need more structure, we](#)  
178 [can move to option 4.](#)

179 [Option 4 is recommended now!](#)

### 180 **9.1 Into What Namespace Do Extensions Go**

181 Extensions (by users) go into user-defined namespaces outside of UBL.

## 182 **10 Options: Module Structure**

183 This section describes options for decomposing schema definitions into modules, where  
184 modules are typically represented as operating system files. For XML Schemas, each file  
185 contains one schema document instance. A more general definition of “module” is as  
186 follows:

187 **Definition:** A Module is a <xsd:schema> document instance. In the UBL deliverable,  
188 each module is written to one operating system file. But in database storage (either  
189 RDBMS or XML native), a module would be recognized as an XML document instance.

190 The following options for module decomposition have been identified:

### 191 ***10.1 Option 1: One Module Per Namespace***

192 This is the option used in the Op70 UBL deliverable. It is the simplest rule to apply and  
193 works reasonably well for the size and scope of the Op70 deliverable. However, it may  
194 not scale to a more mature library of several hundred reusable type definitions. The  
195 scalability concern is not due to technical issues, but due to difficulty of human users  
196 working with one very large file. Tool support will help to mitigate this problem, but  
197 even then some kind of logical modularity would be useful.

### 198 ***10.2 Option 2: One Module Per Object Class***

199 This option would gather together all of the qualified variations of BIEs for each object  
200 class, as implemented by schema type definitions and their associated global elements. So  
201 BuyerParty, SellerParty, and so on would appear in one module.

202 A master schema must include all modules for a given namespace. Users of a namespace  
203 library would not import the individual modules, but only the master schema.

204 The primary motivation for this rule is to provide an easily automated decomposition  
205 strategy that does not require human intervention when generating schemas from a model  
206 or component repository.

207 A downside of this option is that the type definitions in a module do not include any  
208 definitions for closely related content element definitions.

### 209 ***10.3 Option 3: Modules based on Human Judgment of 210 Related Functionality of Type Definitions***

211 This option would gather together related type definitions based on functional similarity.  
212 For example, HazardousItem and its related child element content definitions would be  
213 collected in one module.

214 This might require substantial human analysis to determine the best decomposition of a  
215 namespace into modules. In particular, when leaf schema types (e.g. CountryType) are  
216 used by several modules, those shared types cannot be duplicated in functional modules.

217 **10.4 Option 4: One Module per Type Definition**

218 This is essentially the rule used for creating xCBL modules. Use of the schema files is  
219 only practical when they are opened in a schema design tool. A user would open the  
220 master schema, which must include several hundred small schema files.

221

222 **11 Recommendation: Module Structure**

223 This section describes the mapping of namespaces (as discussed in section 8 *Options:*  
224 *Namespace Structure*) onto XSD files. A namespace contains type definitions and  
225 element declarations. Any file containing type definitions and element declarations is  
226 called a SchemaModule.

227 Every namespace has a special SchemaModule, a RootSchema. Other namespaces  
228 dependent upon type definitions or element declaration defined in that namespace import  
229 the RootSchema and only the RootSchema.

230 If a namespace is small enough then it can be completely specified within the  
231 RootSchema. For larger namespaces, more SchemaModules may be defined – call these  
232 InternalModules. The RootSchema for that namespace then include those  
233 InternalModules.

234 This structure provides encapsulation of namespace implementations. To recap:

235 **Import Rule:** A namespace “A” dependent upon type definitions or element declaration  
236 defined in another namespace “B” imports B’s RootSchema. “A” never imports other  
237 (internal) schema modules of “B”.

238

239 **Include Rule:** The only place XSD “include” is used is within a RootSchema. When a  
240 namespace gets large, its type definitions and element declarations may be split into  
241 multiple SchemaModules (called InternalModules) and included by the RootSchema for  
242 that namespace.

243 The import rule presents a namespace as an indivisible grouping of types. A “piece” of a  
244 namespace can never be used without all it’s pieces. It is therefore important to strive to  
245 define namespaces that are minimal and orthogonal.

246 **Spin out minimal and orthogonal a bit more.**

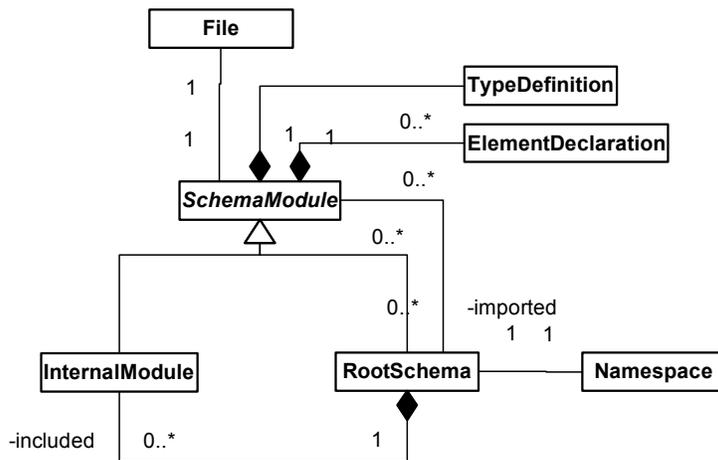
247 It is not enough that a namespace be minimal in terms of its intrinsic size, but also in  
248 terms of the closure of all other namespaces it imports. By closure we mean namespaces  
249 it imports, and namespaces they import, and so on.

250 One good way to foster minimal namespaces is to dictate that there be no circular  
 251 dependencies between them. The same statement can be made for SchemaModules.  
 252 This rule has been applied successfully in many large systems<sup>2</sup>.

253 **(No) Circular Dependency Rule:** There are no circular dependencies between  
 254 SchemaModules. By extension, there are no circular dependencies between namespaces.  
 255 This rule is not limited to *direct* dependencies – transitive dependencies must be taken  
 256 into account.

257 Here is a depiction of the component structure we’ve described so far. This is a UML  
 258 Static Structure Diagram. It uses classes and associations to depict the various concepts  
 259 we’ve been discussing:

260



261

262

263 You can see that there are two kinds of schema module: RootSchema and  
 264 “InternalModule”. A RootSchema may have zero or more InternalModules that it  
 265 includes. Any SchemaModule, be it a RootSchema or an InternalModule may import  
 266 other RootSchemas.

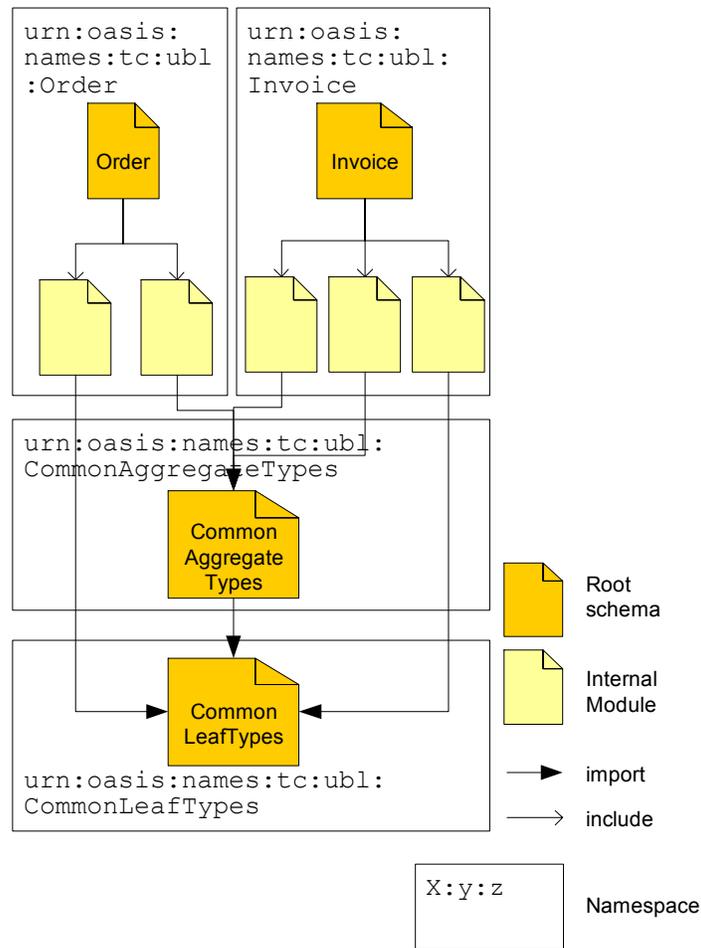
267 The diagram shows the 1-1 correspondence between RootSchemas and namespaces. It  
 268 also shows the 1-1 correspondence between files and SchemaModules. A  
 269 SchemaModule consists of type definitions and element declarations.

270 The diagram unfortunately fails to express the (No) Circular Dependency Rule.

271 Another way to visualize the structure is by example. The following informal diagram  
 272 depicts instances of the various classes from the previous diagram.

---

<sup>2</sup> For example [LARGE-SCALE] introduces the concept of “levelization” as an organizing principal for very large C++ systems. Those systems, due to the nature of the language often have an explosion of type definitions (due to the presence of parameterized types). As a result, solutions to the management of type systems in C++ could be viewed as exemplary for our purposes.



273 The preceding diagram shows how the order and invoice RootSchemas import the  
 274 “CommonAggregateTypes” and “CommonLeaf Types” RootSchemas. It also shows how  
 275 e.g. the order RootSchema includes various InternalModules – modules local to that  
 276 namespace. The clear boxes show how the various SchemaModules are grouped into  
 277 namespaces.

## 278 **11.1 Message Types**

279 If preferring type definitions over global element definitions is good, why not take it to  
 280 the extreme [NDR-MSG-70]. **The type of the root element of a UBL document**  
 281 **(message) is a global type** (not an anonymous type).

## 282 **11.2 Number of Message Types**

283 In some cases, various actions in the protocol (create vs. delete) will have totally different  
 284 document structure requirements. But in some cases (create vs. update), the content might  
 285 be identical. However, we still think we should design in favor of more document types  
 286 rather than less, e.g. one for each transmission (a la RosettaNet). It avoids confusion on  
 287 the part of developers to have a separate document type for each thing. We might then  
 288 decide to optimize some of them by merging them together.

289 **12 Options: Versioning**

290 [XFRNT-VER] does a great job of laying out the problem and solution space for schema  
291 versioning as it is traditionally practiced. The options presented in that document are not  
292 really disjoint rather they are building blocks. If you look at the recommendations in that  
293 document, you will see that the options are used in concert.

294 **12.1 Option XF-1: Change the (internal) schema**  
295 **“version” attribute**

296 **12.2 Option XF-2: Create a “schemaVersion” attribute**  
297 **on the root element**

298 **12.2.1 Usage A: Conformance enforced by validator**

299 **12.2.2 Usage B: Conformance enforced by an extra processing**  
300 **pass**

301 **12.3 Option XF-3: Change the schema’s target**  
302 **namespace**

303 **12.4 Option XF-4: Change the name/location of the**  
304 **schema**

305 **12.5 Option 5: Schema Version as Context Classifier**

306 In [NDR-MSG-13] the point was made that schema version might just be another context  
307 classifier.

308 **13 Recommendations: Versioning**

309 The following table summarizes the tradeoffs between the options.

	Pro	Con
Option XF-1: Change the (internal) schema “version” attribute		Not enforced by validator
Option XF-2-A: Create a “schemaVersion” attribute on the root element -- Conformance enforced by validator		Conformance requires exact match on version string
Option XF-2-B: Create a “schemaVersion” attribute		Extra processing step.

on the root element -- Conformance enforced by an extra processing pass		
Option XF-3: Change the schema's target namespace	Schema validation ensures that an instance conforms to its declared schema. There are never two (different) schemas with the same namespace URI.	With this approach, instance documents will not validate until they are changed to designate the new targetNamespace. However, one does not want to force all instance documents to change, even if the change to the schema is really minor and would not impact an instance.  +Include problems.
Option XF-4: Change the name/location of the schema		Ugh!
Option 5: Schema Version as Context Classifier	Leverages the context machinery	Requires the context machinery

310

311 We will use Option XF3 as a starting point for UBL. A UBL namespace URI is divided  
312 into two parts, one that describes the **purpose** of the namespace and another that captures  
313 **version** information.

314 The version information will in turn be divided into *major* and *minor* fields. For  
315 example, the namespace URI for the Invoice domain has this form:

316 `urn:oasis:names:tc:ubl:schema:Invoice:major-version:minor-version`

317 The *major-version* field is "1" for the first release of a namespace. Subsequent major  
318 releases increment the value by 1. For example, the first namespace URI for the first  
319 major release of the Invoice domain has the form:

320 `urn:oasis:names:tc:ubl:schema:Invoice:1:0`

321 The second major release will have a URI of the form:

322 `urn:oasis:names:tc:ubl:schema:Invoice:2:0`

323 The distinguished value "0" (zero) is used in the *minor-version* position when defining a  
324 new major version. In general, the namespace URI for every major release of the Invoice  
325 domain has the form:

326 `urn:oasis:names:tc:ubl:schema:Invoice:major-number:0`

327 Subsequent minor releases begin with *minor-version* 1. For example, the namespace URI  
328 for the first minor release of the Invoice domain has this form:

329 `urn:oasis:names:tc:ubl:schema:Invoice:major-number:1`

330 In UBL, the major-version field of a namespace URI must be changed in a release that  
331 breaks compatibility with the previous release of that namespace. If a change does not  
332 break compatibility then only the minor version need change. Regardless, at a minimum  
333 any change to any schema module constituting the namespace necessitates some change  
334 to the namespace URI. Said another way, **once a namespace URI is published by UBL**  
335 **it must never change.**

336 This approach yields non-obvious, yet beneficial effects when the interdependencies of  
337 namespaces are considered. UBL is composed of a number of interdependent  
338 namespaces. For instance, namespaces whose URI's start with  
339 `urn:oasis:names:tc:ubl:schema:Invoice:*` are dependent upon the common leaf  
340 and aggregate namespaces, whose URI's have the form  
341 `urn:oasis:names:tc:ubl:schema:CommonLeafTypes:*` and  
342 `urn:oasis:names:tc:ubl:schema:CommonAggregateTypes:*` respectively. If either  
343 of the common namespaces changes then its namespace URI must change. If its  
344 namespace URI changes then any schema that imports the *new version* of the namespace  
345 must also change (to update the namespace declaration). And if the importing schema  
346 changes then its namespace URI in turn must change. The outcome is twofold:

- 347 • There is never ambiguity at the point of reference. A dependent  
348 schema imports precisely the version of the namespace that is needed.  
349 The dependent never needs to account for the possibility that the  
350 imported namespace can change.
- 351 • When a dependent is upgraded to import a new version of a schema  
352 the dependent's version (in its namespace URI) must change.

353 The question now arises: what is meant by “major” versus “minor”. What kind of change  
354 may a minor version introduce? When is it necessary to incur a new major version?  
355 Why are the answers to these questions even interesting?

356 To answer these questions you must start by understanding that UBL's use of major and  
357 minor version number borrows from a long tradition software tradition. In that tradition,  
358 a minor version declared it's “compatibility” with previous minor versions (of the same  
359 major version).

360 Since this sort of versioning scheme was applied to libraries, applications and even whole  
361 operating systems, the definition of the term “compatibility” in those various contexts  
362 necessarily varied widely.

363 Its historical use in shared libraries probably comes closest to the intended UBL use. A  
364 new release of a library (namespace) must specify a new major version number if it  
365 breaks compatibility with the previous version of the library (namespace). In the case of  
366 object libraries examples of breaking compatibility were 1) calling interface changed or  
367 2) behavior (semantics) of interface changed.

368 Implicit in this major/minor scheme is that there is some benefit in breaking the version  
369 information into two pieces. The benefit in the traditional shared library paradigm is that  
370 objects dependent upon those shared libraries could still function properly with  
371 subsequent minor releases. Those minor releases might add new functionality or repair  
372 defects – but they wouldn't break the “contract” identified by the major version number.

373 The level of formal specification of this “contract” has varied in historical practice  
374 ranging from informal and undocumented to human-readable interface specification.  
375 UBL leverages XML schema itself as the means to capture this contract. Here’s how it  
376 works...

377 A minor revision to a major release (of a namespace) *imports* the schema module for the  
378 major release. For instance, the schema module defining:

379 `urn:oasis:names:tc:ubl:schema:Invoice:1:2`

380 *Must* import the namespace:

381 `urn:oasis:names:tc:ubl:schema:Invoice:1:1`

382 The 1:2 revision may define new complex types by extending or restricting 1:1 ones. It  
383 may define brand new complex types and elements by composition. It must not use the  
384 XSD redefine element to change the definition of a type or element in the 1:1 version.

385 The opportunity exists in the 1:2 version to rename derived types. For instance if 1:1  
386 defines Address and 1:2 specializes Address it would be possible to give the derived  
387 Address a new name, e.g. NewAddress. This is not required since namespace  
388 qualification suffices to distinguish the two distinct types. **The minor revision may give  
389 a derived type a new name only if the derived type represents a semantic distinct  
390 from that of the base type.**

391 For a particular namespace, the minor versions of a major version form a linearly-linked  
392 family. Each successive minor version imports the schema module of the preceding  
393 minor version. The process is bootstrapped by the first minor version importing the  
394 namespace defining the major version of interest. E.g.

395 `urn:oasis:names:tc:ubl:schema:Invoice:1:2 imports`

396 `urn:oasis:names:tc:ubl:schema:Invoice:1:1 which imports`

397 `urn:oasis:names:tc:ubl:schema:Invoice:1:0.`

398 The outcome of this usage of XSD import is that schema validation enforces these  
399 constraints:

- 400 1. forward compatibility of instances: an instance document valid in version M:m  
401 will be valid in any version M:m+n.
- 402 2. backward compatibility of reused components: an instance document that is valid  
403 in version M:m may contain constructs defined in M:m-n. Processing logic  
404 implemented in terms of version M:m-n will process those constructs properly  
405 since those constructs are valid with respect to version M:m-n.
- 406 3. backward incompatibility of new constructs: new constructs defined in version  
407 M:m will not be valid in M:m-n therefore processing logic would not be expected  
408 to operate on them.
- 409 4. potential backward compatibility of extended constructs: Extensions (of complex  
410 types) defined in M:m are *valid* in M:m-n however, processing logic implemented  
411 in terms of M:m-n will not be aware of extension elements. Care must be taken in  
412 the construction of processing logic to maximize the potential for compatible  
413 extension. In particular, processing logic that copies element content should do so  
414 in such a way that extension elements will be copied too.

415 When the changes to a namespace are such that it doesn't fit conveniently into this  
416 scheme (of importing the previous minor version namespace), a new major version is  
417 created. A new major version does not import previous version namespaces, nor does it  
418 make any representation as to compatibility with old versions. The purpose of the major  
419 version is to free the UBL designers to make significant, incompatible changes to the  
420 library.

421 It bears stating explicitly again that UBL is composed of a number of interdependent  
422 namespaces. It is not a single monolithic component. While it is expected that UBL  
423 releases will be assigned version identifiers of some sort e.g. UBL 1, UBL 2, this should  
424 not be confused with the versioning of the UBL namespaces discussed in this section. It  
425 would be perfectly reasonable, for example, for a release called "UBL version 2" to  
426 contain namespaces with URI's whose major version is not 2. Namespace versioning as  
427 described here is a fine-grained, technical mechanism for declaring and enforcing  
428 compatibility between interdependent namespaces over long periods of time (years).

429

## 430 14 Definitions

431 Backward compatibility – TBD.

432 **BIE** – Business Information Entity. A description of a business concept. Represented as an XML schema  
433 by a *root schema*.

434 **extension** a.k.a. customization – specification of new BIE's with well-defined, enforced relationships to old  
435 BIE's. Relationship types include: restriction, extension. In some cases processing logic will need to treat  
436 the base and the extension as the same, in other cases it will need to distinguish between them.

437 Forward compatibility – TBD

438 **Namespace** – a name that scopes a related group of XML type definitions.

439 **processing logic** – software logic that operates on BIE instances to achieve some business function

440 **root schema** – A *schema module* that directly, or via inclusion of other schema modules, defines all types  
441 for a particular namespace. This is the XML Schema representation of a BIE. (Compare that definition,  
442 with the one we came up with last week in Menlo Park: *A schema document corresponding to a single*  
443 *namespace, which is likely to pull in (by including or importing) schema modules. Issue: Should a root*  
444 *schema always pull in the "meat" of the definitions for that namespace, regardless of how small it is?)*

445 **schema document** – as defined by the XSD specification – per that specification, a schema document  
446 defines types into exactly one namespace, the target namespace.

447 **schema module** – A *schema document*. A schema module need not define all types in a  
448 particular namespace. Contrast with *root schema*. (Compare that definition, with last  
449 week's: *A "schema document" (as defined by the XSD spec) that is intended to be taken*  
450 *in combination with other such schema documents to be used.*)

451 **versioning** – reification of revisions to BIE's in order to support coexistence in a system,  
452 of two or more revisions of a BIE.

## 453 15 References

LARGE-SCALE	<i>Large Scale C++ Software Design,</i>	
-------------	---	--

	John Lakos, 1996, Addison-Wesley.	
NAMESPACE	<i>Namespaces in XML</i>	<a href="http://www.w3.org/TR/REC-xml-names/">http://www.w3.org/TR/REC-xml-names/</a>
NDR-MSG-13	<i>schema version as context classifier</i> , Burcham, Bill; Maler, Eve; a post to the UBL-NDR mailing list.	<a href="http://lists.oasis-open.org/archives/ubl-ndrsc/200111/msg00013.html">http://lists.oasis-open.org/archives/ubl-ndrsc/200111/msg00013.html</a>
NDR-MSG-70	<i>Fwd: Straw Man on Namespaces, Schema Module Architecture, etc.</i> , Rawlins, Mike; a post to the UBL-NDR mailing list.	<a href="http://lists.oasis-open.org/archives/ubl-ndrsc/200111/msg00070.html">http://lists.oasis-open.org/archives/ubl-ndrsc/200111/msg00070.html</a>
NDR-MSG-88	<i>Fwd: Straw Man on Namespaces, Schema Module Architecture, etc.</i> , Probert, Sue; Maler, Eve.; a post to the UBL-NDR mailing list.	<a href="http://lists.oasis-open.org/archives/ubl-ndrsc/200111/msg00088.html">http://lists.oasis-open.org/archives/ubl-ndrsc/200111/msg00088.html</a>
OASIS-URN-NS	IETF RFC 3121 <i>A URN Namespace for OASIS</i>	<a href="http://www.faqs.org/rfcs/rfc3121.html">http://www.faqs.org/rfcs/rfc3121.html</a>
SCHEMA-PRIM	<i>XML Schema Part 0: Primer</i>	<a href="http://www.w3.org/TR/xmlschema-0/">http://www.w3.org/TR/xmlschema-0/</a>
SEVEN-TWO	<i>The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information</i> , George A. Miller, Psychological Review, 63, 81-97.	<a href="http://psychclassics.yorku.ca/Miller/">http://psychclassics.yorku.ca/Miller/</a>
XFRNT-VER	<i>XML Schema Versioning</i> , MITRE Corporation and xml-dev list group members.	<a href="http://www.xfront.com/Versioning.pdf">http://www.xfront.com/Versioning.pdf</a>
XML-URI-LIST	<i>XML-URI List</i> at w3.org	<a href="http://lists.w3.org/Archives/Public/xml-uri/">http://lists.w3.org/Archives/Public/xml-uri/</a>

454

455