Regular Language Description for XML (RELAX): RELAX Core

1 Scope

This Technical Report specifies mechanisms for formally specifying the syntax of XML-based languages. For example, the syntax of XHTML 1.0 can be specified in RELAX.

Compared with DTDs, RELAX provides the following advantages:

- Specification in RELAX uses XML instance (i.e., document) syntax,
- RELAX provides rich datatypes, and
- RELAX is namespace-aware.

The RELAX specification consists of two parts, *RELAX Core* and *RELAX Namespace*. This Technical Report specifies *RELAX Core,* which may be used to describe markup languages containing a single XML namespace. Part 2 of this Technical Report specifies *RELAX Namespace*, which may be used to describe markup languages containing more than a single XML namespace, consisting of more than one *RELAX Core* document.

Given a sequence of elements, a software module called the *RELAX Core* processor compares it against a specification in *RELAX Core* and reports the result. The *RELAX Core* processor can be directly invoked by the user, and can also be invoked by another software module called the *RELAX Namespace* processor.

This Technical Report also specifies a subset of *RELAX Core*, which is restricted to DTD features plus datatypes. This subset is very easy to implement, and with the exception of datatype information, conversion between this subset and XML DTDs results in no information loss.

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of this Technical Report.

W3C (World Wide Web Consortium), Extensible Markup Language (XML) 1.0, W3C Recommendation, http://www.w3.org/TR/REC-xml, 1998

W3C (World Wide Web Consortium), Name Spaces in XML, W3C Recommendation, http://www.w3.org/TR/REC-xml-names, 1999

W3C (World Wide Web Consortium), XML Information Set, W3C Working Draft, http://www.w3.org/TR/xml-infoset

W3C (World Wide Web Consortium), XML Schema Part 2, W3C Working Draft, http://www.w3.org/TR/xmlschema-2

IETF (Internet Engineering Task Force). RFC2396: Uniform Resource Identifiers (URI): Generic Syntax, 1998.

3 Terms and definitions

3.1 XML 1.0

For the purposes of this Technical Report, the following terms and definitions given in XML 1.0 apply.

a) start tag

b) end tag

c) empty-element tag

d) attribute

e) attribute name

f) content

g) content model

h) attribute-list declaration

i) DTD

j) XML processor

k) validity

l) validating processor

m) non-validating processor

n) whitespace

o) child

p) parameter entity

q) match

NOTE On top of those meanings given in XML 1.0, “match” has another meaning (see 5.8).

3.2 Name Spaces in XML

For the purposes of this Technical Report, the following terms and definitions given in “Name Spaces in XML” apply.

a) namespace

b) namespace name

3.3 XML Schema Part 2

For the purposes of this Technical Report, the following terms and definitions given in "XML Schema Part 2" apply.

a) lexical representation

b) facet

c) datatype

d) built-in datatype

3.4 XML Information Set

For the purposes of this Technical Report, the following terms and definitions given in “XML Information Set” apply.

a) information set

b) document information item

c) element information item

d) property

e) core property

f) reference to skipped entity information item

g) entity information item

h) notation information item

3.5 Definitions specific to *RELAX Core*

3.5.1
tag name
names in start tags, end tags, and empty-element tags (generic identifiers in ISO 8879)

3.5.2
hedge
ordered sequences of elements and character data

4 Notations

This Technical Report uses DTD in order to specify the syntax of RELAX modules. However, since DTDs provide no support for XML namespaces, this Technical Report only uses some of the constructs possible in DTDs.

To specify permissible contents of elements, this Technical Report uses content models, which match the non-terminal symbol `contentspec` in XML 1.0.

EXAMPLE 1 The following content model specifies that an element is constrained to a sequence beginning with a **frontmatter** element followed by a **body** element, and finally an optional **backmatter** element

```
(frontmatter, body, backmatter*)
```

To specify permissible attributes of elements, this Technical Report uses fragments of attribute-list declarations, which match the non-terminal symbol `AttDef` of XML 1.0

EXAMPLE 2 The following attribute-list fragment specifies that an element has an optional attribute **class** and that any character string can be used as the attribute value.

```
class CDATA #IMPLIED
```

5 Basic concepts

5.1 Design principles

The design principles of *RELAX Core* are:

a) *RELAX Core* shall be simple and powerful.

b) The design shall be prepared quickly.

c) The design shall be formal and concise.

d) It shall be possible to implement *RELAX Core* using existing XML document APIs (e.g., SAX and DOM).

e) *RELAX Core* shall be upward-compatible with DTDs.

f) *RELAX Core* shall have a subset such that conversion to and from DTDs loses no information except datatype information.

g) Datatypes of *RELAX Core* shall be compatible with those in XML Schema Part 2.

5.2 Instances, grammars, and meta grammars

5.2.1 Instances

A document information item is said to be an *instance*. When an instance satisfies conditions represented by a RELAX grammar, the instance is said to *comply* or be *compliant with* the RELAX grammar. If there is no possibility of confusion, the instance may be considered compliant without mention of the RELAX grammar.

NOTE A valid document as defined in XML 1.0 (to be precise, a document information item represented by this document) need not be compliant with a RELAX grammar; an instance compliant with a RELAX grammar (to be precise, documents representing this instance) need not be valid.

5.2.2 RELAX grammars

A document information item that conforms to *RELAX Namespace* is said to be a *RELAX grammar*.

5.2.3 RELAX meta grammars

The *RELAX meta grammar* is a RELAX grammar specifying the syntax of RELAX. Any RELAX grammar is compliant with the RELAX meta grammar.

5.3 Modules and grammars

A document information item that conforms to *RELAX Core* is said to be a *RELAX module*. A RELAX module addresses elements in a single namespace as well as their attributes and contents.

Since a single-namespace RELAX grammar references only a single module, the module provides the complete grammar definition.

A multiple-namespace RELAX grammar references and combines multiple modules. Such a RELAX grammar is introduced in *RELAX Namespace*.

5.4 Islands and instances

A multi-namespace instance is compared against a RELAX grammar comprising multiple modules. Such an instance is first decomposed into multiple *islands*, each of which is a single-namespace hedge. Each island is then compared against a single RELAX module (Figure 1).

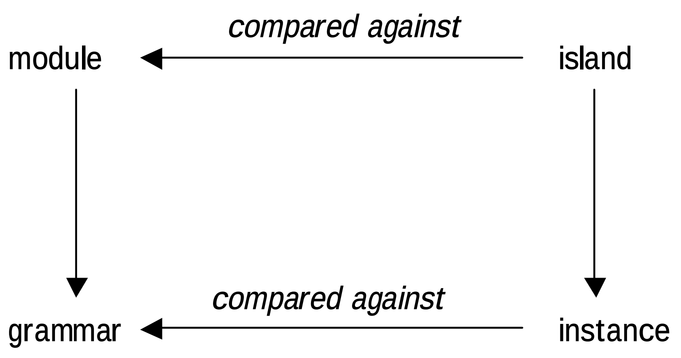

Figure 1 — The relationship between modules/grammars and islands/instances

A single-namespace instance is already an island, and thus need not be further decomposed.

5.5 Behaviour of the *RELAX Core* processor

The *RELAX Core* processor is a software module that, given an island and a RELAX module, compares the island against the RELAX module in order to determine if the island is compliant with the RELAX module.

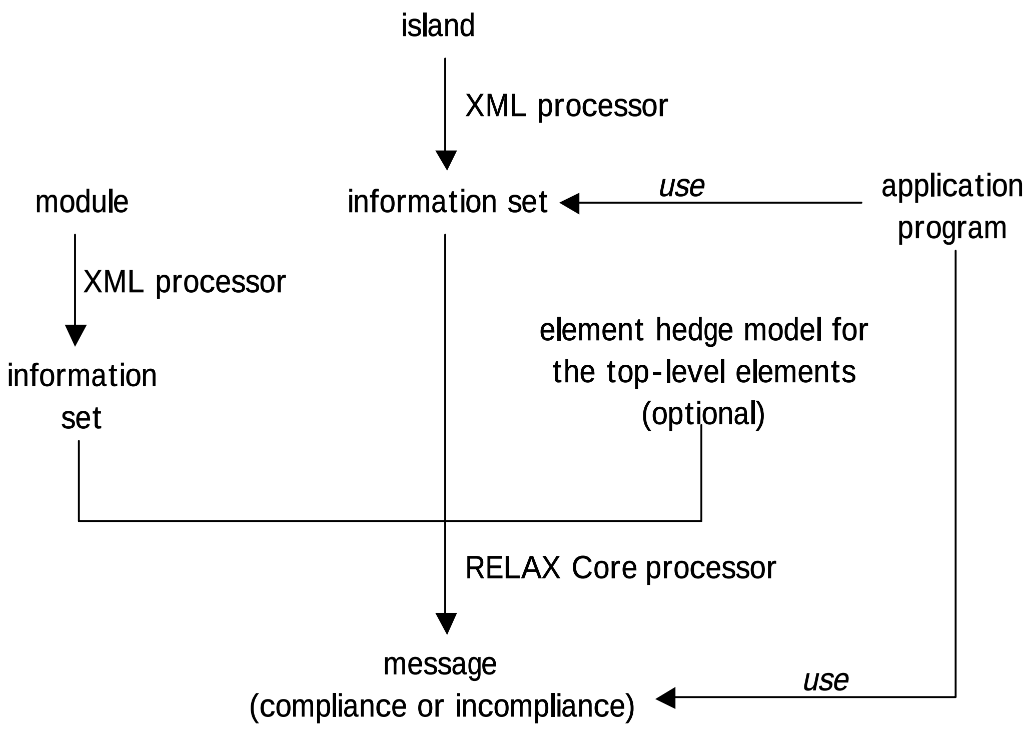

Figure 2 — The *RELAX Core* processor, the XML processor, and application programs

The *RELAX Core* processor shall receive islands and RELAX modules as information sets from the XML processor. The *RELAX Core* processor shall use core properties of information items in the information sets and shall not use other properties.

NOTE Implementations of the *RELAX Core* processor receive information sets via APIs such as SAX or DOM.

The *RELAX Core* processor may also receive a hedge model that constrains the top-level elements of islands.

After comparison, the *RELAX Core* processor shall output a message that the island is compliant or a message that it is not. The *RELAX Core* processor may output other messages.

Other than such messages, the *RELAX Core* processor shall have no outputs. Application programs shall receive information sets from the XML processor and may receive messages from the *RELAX Core* processor.

Both validating processors and non-validating processors may be used by the *RELAX Core* processor.

When the *RELAX Core* processor receives references to skipped entity information items, it shall output a message, at user option, and may stop normal processing.

5.6 Datatypes

RELAX Core uses the built-in datatypes of XML Schema Part 2. Datatypes can be used as conditions on attributes or used as hedge models.

Datatypes in *RELAX Core* represent sets of strings. Given a string and a datatype, it is possible to determine if the string is contained by the set of strings represented by that datatype.

EXAMPLE 1 The datatype named **integer** defines a set of strings representing integers. It is possible to determine whether or not a string represents an integer.

References to datatypes may have additional conditions called *facets*.

EXAMPLE 2 A reference to the **integer** datatype may have facets, which may specify that the value should be equal to or greater than 10, and that it should be equal to or less than 20.

The *RELAX Core* processor does not convert character strings to data (e.g., conversion of the string "1" to the integer 1). Such conversion is left to application programs.

NOTE Typically, application programs merely invoke conversion libraries.

5.7 Roles and clauses

In *RELAX Core*, conditions on tag names and attributes are captured by *roles* and *clauses*. A role is a name, and is described by a clause. A clause does not have a name.

A clause is either **tag** or **attPool**. **tag** has a condition on tag names, while **attPool** does not. When **tag** or **attPool** contains conditions on an attribute, this attribute is said to be *declared* by the **tag** or **attPool**.

There shall be at most one clause per role.

A clause may reference another clause via a role. A clause shall not directly or indirectly reference itself. A referenced clause shall be described by **attPool**; it shall not be described by **tag**.

A clause shall not directly or indirectly (via other clause) reference another clause more than once.

A clause shall not directly or indirectly (via other clause) declare an attribute more than once.

5.8 Production rules, labels, and hedge models

5.8.1 General

In *RELAX Core*, conditions on element structures are captured by *labels* and *production rules*. A label is a name, and is described by a production rule. A production rule shall not have a name.

A production rule is either **elementRule** or **hedgeRule**. **elementRule** is a triplet of a label, role, and hedge model. **hedgeRule** is a pair, consisting of a label and a hedge model.

Roles referenced by **elementRule** shall be described by **tag** clauses. **elementRule** shall not reference roles described by **attPool** clauses.

More than one **elementRule** may share a label, and more than one **hedgeRule** may share a label. However, **elementRule** and **hedgeRule** shall not share a label.

More than one **elementRule** may share a role.

NOTE Regular grammars comprise production rules and generate sets of strings. The left-hand side of a production rule is a non-terminal symbol, and the right-hand side is either a terminal symbol, a terminal symbol followed by a non-terminal symbol, or a non-terminal symbol. *RELAX Core* is an extension of regular grammars such that sets of logical structures are generated. A label corresponds to a non-terminal symbol in the left-hand side, and a role corresponds to a terminal symbol. Hedge models extend non-terminal symbols in the right-hand side so that **elementRule** addresses tree structures rather than strings.

Permissible contents of elements are described by *hedge models*. A hedge model shall be either an element hedge model, mixed hedge model, or datatype reference.

When more than one **elementRule** shares a label and role together, one of the following conditions hold:

a) all of them have element hedge models,

b) all of them have mixed hedge models, or

c) all of them reference the same datatype, possibly having different facets.

5.8.2 Element hedge models

An element hedge model generates a regular set of label sequences. Any label sequence in this set, possibly prepended, interspersed with, or followed by whitespace characters, is said to *match* this element hedge model.

5.8.3 Mixed hedge models

A mixed hedge model is an element hedge model with the **mixed** wrapper. Any label sequence in the set generated by the wrapped element hedge model, possibly prepended, interspersed with, or followed by arbitrary characters, is said to *match* this mixed hedge model. Note that not only whitespace characters but also non-whitespace characters are permitted.

5.8.4 Datatype references

A datatype reference specifies a datatype name. It may further specify additional conditions called facets. A sequence of characters *matches* a datatype reference if the character sequence belongs to the referenced datatype and satisifies the accompanying facets, if any.

5.9 Taxonomy and occurrences of names

Five types of names are used in *RELAX Core*. They are datatype names, tag names, attribute names, roles, and labels. Names of different types do not collide. For example, although there is a datatype called **integer**, one may use "integer" as a tag name or label.

The following table shows which types of names appear in instances and where in RELAX modules they appear.

Table 1 — Types of names and their occurrences in instances and RELAX modules

Types of names	In Instances	In RELAX Modules
datatype names	do not occur	occur as conditions on attributes or hedge models (datatype references)
tag names	occur	occur as part of clauses
attribute names	occur	occur as part of clauses
roles	do not occur	occur in clauses (description of roles) occur in clauses (references to roles)
labels	do not occur	occur in production rules (description of labels) occur in production rules (reference to labels)

6 Module Constructs

6.1 module

module represents an entire module. This element provides management information about the module.

module has the **moduleVersion** attribute the ***relaxCore*Version** attribute, and the **targetNamespace** attribute.

```
moduleVersion CDATA #IMPLIED
relaxCoreVersion CDATA #REQUIRED
targetNamespace CDATA #IMPLIED
```

The version of this module is indicated by the **moduleVersion** attribute.

The version of *RELAX Core* is indicated by the **relaxCoreVersion** attribute. The version number "1.0" shall be used to indicate conformance to Version 1.0 of *RELAX Core*; it is an error for a document to use the value "1.0" if it does not conform to Version 1.0 of *RELAX Core*.

NOTE It is the intent of the editing committee to give later versions of this specification numbers other than "1.0", but this intent does not indicate a commitment to produce any future versions of *RELAX Core*, nor if any are produced, to use any particular numbering scheme. Since future versions are not ruled out, this construct is provided as a means to allow the possibility of automatic version recognition, should it become necessary.

The *RELAX Core* processor may signal an error if it receives documents labelled with versions that it does not support. It may continue or abort normal processing.

The **targetNamespace** attribute specifies the namespace to which elements described by this module belong. If this attribute is omitted, the value "" is assumed.

Given that all elements in the module itself belong to the namespace "http://www.xml.gr.jp/xmlns/relaxCore", the **module** element shall declare this namespace.

Every element in a module shall belong to the namespace "http://www.xml.gr.jp/xmlns/*RELAX Core*". Thus, the module element shall declare this namespace.

The following content model describes the permissible content of **module** elements:

```
(annotation?, interface?,
  (tag | attPool | elementRule | hedgeRule | div | include )*)
```

EXAMPLE An example **module** element is shown below. Child elements are omitted for clarity.

```
<module
   moduleVersion="1.2"
   RELAX CoreVersion="1.0"
   xmlns="http://www.xml.gr.jp/xmlns/RELAX Core">
   ...
</module>
```

6.2 interface

interface provides interface information between the module and RELAX grammars. If a RELAX grammar has only a single namespace, **interface** provides information about the permissible root element of instances.

interface has no attributes.

The following content model describes the permissible content of **interface** elements:

```
(annotation?, (export | div)*)
```

6.3 export

export indicates which information can be referenced from RELAX grammars.

export has the **label** attribute.

```
     label NMTOKEN #REQUIRED
```

The **label** attribute exposes a label described by **elementRule** to RELAX grammars. If a RELAX grammar has only a single namespace, this attribute shows that the root of instances may have a specified label.

The following content model describes the permissible content of **export** elements:

```
      (annotation?)
```

EXAMPLE An example of **export** is shown below.

```
     <export label="doc"/>
```

6.4 tag

tag specifies the condition that elements play a specified role, by combining a condition on tag names, conditions on attribute values, and references to other roles.

tag has the role attribute and the name attribute.

```
     role NMTOKEN #IMPLIED
     name NMTOKEN #IMPLIED
```

The **name** attribute specifies the tag name. The **role** attribute specifies which role is described by the **tag**.

When the **tag** is not a child element of **elementRule**, the **name** attribute shall be specified but the **role** attribute need not be specified. If the **role** attribute is not specified, its value is assumed to be the same as the **name** attribute.

When the **tag** is a chid element of **elementRule**, the **name** attribute need not be specified and the **role** attribute shall not be specified. An appropriate role which does not collide with other roles is generated as the value of the **role** attribute by the *RELAX Core* processor. If the **name** attribute is not specified, it is assumed to be the same as the **label** attribute of the parent **elementRule**.

The following content model describes the permissible content of **tag** elements. Subordinate **ref** shall specify the **role** attribute.

```
     (annotation?, ref*, attribute*)
```

An element *e* play the **role** specified by the **role** attribute if the following three conditions hold:

a) The tag name of *e* matches the value of the **name** attribute.

b) Each of the conditions (expressed by the subordinate **atttribute** elements) on attributes is satisfied by some attribute of *e*.

c) *e* plays all roles referenced by the subordinate **ref** elements.

Observe that *e* satisfies these conditions even if it has attributes not declared by the **tag.**

EXAMPLE An example of **tag** is shown below. The **bar1** role is referenced by a subordinate **ref** element.

```
     <tag name="foo" role="bar">
       <ref role="bar1"/>
     </tag>
```

6.5 attPool

attPool specifies the condition that elements play a specified role, by combining conditions on attribute values, and references to other roles.

attPool has the **role** attribute.

```
     role NMTOKEN #REQUIRED
```

The **role** attribute specifies which role is described by the **attPool**.

The following content model describes the permissible content of **attPool** elements. Subordinate **ref** shall specify the **role** attribute.

```
(annotation?, ref*, attribute*)
```

An element *e* plays the role specified by the **role** attribute if the following two conditions hold:

a) Each of the conditions (expressed by the subordinate **atttribute** elements) on attributes is satisfied by some attribute of *e*.

b) *e* plays all roles specified by the subordinate **ref** elements.

EXAMPLE An example of **attPool** is shown below. The **bar1** role is referenced by a subordinate ref element.

```
<attPool role="bar">
  <ref role="bar1"/>
</attPool>
```

6.6 ref with the role attribute

ref with the **role** attribute references a role described by **attPool**.

ref with the **role** attribute does not have other attributes.

```
role NMTOKEN #REQUIRED
```

The following content model describes the permissible content of **ref** elements having the **role** attribute:

```
EMPTY
```

Examples of **ref** with the **role** attribute are also contained in the examples of **tag** and **attPool**.

6.7 attribute

attribute describes conditions on attribute names and values. It further indicates whether the attribute is optional.

attribute has the **name** attribute, the **required** attribute, and the **type** attribute. The **name** attribute is mandatory.

```
name NMTOKEN #REQUIRED
required (true) #IMPLIED
type NMTOKEN #IMPLIED
```

The **name** attribute specifies the attribute name. The **required** attribute shows whether this attribute is optional. If "**true**" is specified, this attribute is mandatory.

Multiple **attribute** elements in a **tag** element or the **attPool** elements directly or indirectly referenced by this **tag** shall not specify names matching each other.

The **type** attribute specifies a datatype name. If the **type** attribute is omitted, the built-in datatype **string** is assumed.

The following content model describes the permissible content of **attribute** elements. It is assumed that all element types representing facets are connected by "|" and declared as the value of the `facet` parameter entity.

```
(annotation?, (%facet;)*)
```

The datatype name specified by the **type** attribute and the facets specified by the child elements collectively form a datatype reference. The value of the attribute specified by the **name** attribute is required to match this datatype reference.

EXAMPLE An example of **attribute** is shown below. An **attribute** element is used as a child element of **tag**.

```
<tag name="a" >
  <attribute name="href" type="uriReference"/>
</tag>
```

NOTE In XML 1.0, element types without attributes do not require attribute-list declarations. However, in *RELAX Core*, tag names occurring in instances always require **tag** elements.

6.8 elementRule

elementRule represents a production rule which consists of a triplet of a label, role, and hedge model.

elementRule has the role attribute, the **label** attribute and the **type** attribute. When **elementRule** does not have a subordinate **tag**, the **role** attribute shall be specified and the **label** attribute need not be specified. When the **label** attribute is omitted, it is assumed to have the same value as the **role** attribute. When **elementRule** has a subordinate **tag**, the **label** attribute shall be specified and the **role** attribute shall not be specified.

```
role NMTOKEN #IMPLIED
label NMTOKEN #IMPLIED
type NMTOKEN #IMPLIED
```

The **role** attribute specifies a role. The **label** attribute specifies which label is described by this **elementRule**. The **type** attribute references to a datatype. When this **elementRule** has an element hedge model or mixed hedge model, the **type** attribute shall not be specified.

The following content model describes the permissible content of **elementRule** elements. It is assumed that all element types representing facets are connected by "|" and declared as the value of the `facet` parameter entity. Subordinate **ref** shall specify the **label** attribute.

```
(annotation?, tag?,
  ((ref | hedgeRef | choice | sequence | element | none | empty | mixed)
    |
   (%facet;)*))
```

Handling of subordinate **tag** elements is described in 8.5.

elementRule is said to have an element hedge model if either **ref**, **hedgeRef**, **choice**, **sequence**, **element**, **none** or **empty** is specified as the child element.

elementRule is said to have a mixed hedge model if **mixed** is specified as the child element.

If **elementRule** does not have an element hedge model or mixed hedge model, a datatype reference shall be specified by the **type** attribute. When **elementRule** has a datatype reference, facets may be specified as the content of the **elementRule**. Datatypes and facets are described in Clause 7.

6.9 hedgeRule

hedgeRule represents a production rule which consists of a pair of a label and hedge model.

hedgeRule has the **label** attribute.

```
label NMTOKEN #REQUIRED
```

The **label** attribute specifies which label is described by the **hedgeRule**.

The following content model describes the permissible content of **hedgeRule** elements. Subordinate **ref** shall specify the label attribute.

```
(annotation?,
  (ref | hedgeRef | choice | sequence | element | none | empty))
```

6.10 ref with the label attribute

ref with the **label** attribute represents an element hedge model which references to a label not described by **hedgeRule**.

ref elements with the **label** attribute have the **occurs** attribute.

```
label NMTOKEN #REQUIRED
occurs CDATA #IMPLIED
```

The value of the **occurs** attribute shall be either "*", "+", or "?".

The following content model describes the permissible content of **ref** elements with the **label** attribute:

```
EMPTY
```

Let *l* be the label referenced by the **label** attribute of **ref**. This **ref** shall generate a label sequence made up from one occurrence of *l* only. However, when the **occurs** attribute is specified, "*" shall repeat the sequence zero or more times, "+" shall repeat the sequence one or more times, and "?" shall repeat the sequence zero or one time.

6.11 hedgeRef

hedgeRef represents an element hedge model which references to a label described by some **hedgeRule** element.

hedgeRef has the occurs attribute.

```
label NMTOKEN #REQUIRED
occurs CDATA #IMPLIED
```

The value of the **occurs** attribute shall be either "*", "+", or "?".

The following content model describes the permissible content of **hedgeRef** elements:

```
EMPTY
```

hedgeRef is replaced by element hedge models of those **hedgeRule** elements which describe the label referenced by this **hedgeRef** (more about this, see 8.4).

6.12 squence

sequence represents an element hedge model that concatenates element hedge models.

sequence has the **occurs** attribute. Permissible values and semantics of the **occurs** attribute are the same as in **ref** with the **label** attribute.

```
occurs CDATA #IMPLIED
```

The following content model describes the permissible content of **sequence** elements:

```
(ref | hedgeRef | choice | sequence | element | none | empty)*
```

Suppose that the child elements of **sequence** are $c_1, c_2,\ldots, c_m$. Further suppose that c_1 generates a label sequence $l_{11}, l_{12},\ldots, l_{1n1;}$ c_2 generates a label sequence $l_{21}, l_{22},\ldots, l_{2n2}$; c_3 and the following child elements also generates similar sequences, and c_m generates a label sequence $l_{m1}, l_{m2},\ldots, l_{mnm}$. Then, this **sequence** shall generate $l_{11}, l_{12},\ldots, l_{1n1}, l_{21}, l_{22},\ldots, l_{2n2}, \ldots, l_{m1}, l_{m2},\ldots, l_{mnm}$. However, when the **occurs** attribute is specified, "*" shall repeat the sequence zero or more times, "+" shall repeat the sequence one or more times, and "?" shall repeat the sequence zero or one time.

6.13 choice

choice represents an element hedge model that is a selection from element hedge models.

choice has the **occurs** attribute. Permissible values and semantics of the **occurs** attribute are the same as in **ref** with the **label** attribute.

```
occurs CDATA #IMPLIED
```

The following content model describes the permissible content of **choice** elements:

```
(ref | hedgeRef | choice | sequence | element | none | empty)*
```

Suppose that the child elements of **choice** are $c_1, c_2,\ldots, c_m$. Further suppose that c_1 generates a label sequence $l_{11}, l_{12},\ldots, l_{1n1;}$ c_2 generates a label sequence $l_{21}, l_{22},\ldots, l_{2n2}$; c_3 and the following child elements also generate similar sequences, and c_m generates a label sequence $l_{m1}, l_{m2},\ldots, l_{mnm}$. Then, this **choice** shall generate any of the label sequence $l_{11}, l_{12},\ldots, l_{1n1}$, the label sequence $l_{21}, l_{22},\ldots, l_{2n2}$, ..., or the label sequence $l_{m1}, l_{m2},\ldots, l_{mnm}$. However, when the **occurs** attribute is specified, "*" shall repeat the sequence zero or more times, "+" shall repeat the sequence one or more times, and "?" shall repeat the sequence zero or one time.

6.14 empty

empty represents an element hedge model that matches the empty sequence of labels.

empty has no attributes.

The following content model describes the permissible content of **empty** elements:

```
EMPTY
```

6.15 none

none represents an element hedge model that matches no label sequences.

none has no attributes.

The following content model describes the permissible content of **none** elements:

```
EMPTY
```

6.16 mixed

mixed provides a mixed hedge model.

mixed has no attributes.

The following content model describes the permissible content of **mixed** elements:

```
(ref | hedgeRef | choice | sequence | element | none | empty)
```

Suppose that the element hedge model which is the child of a mixed hedge model generates some label sequence. This label sequence, possibly prepended, intervened, or followed by arbitrary characters, matches the mixed hedge model.

6.17 element

element provides a convenient shorthand which is expanded to **ref**, **tag**, and **elementRule**.

element has the **name** attribute, the **type** attribute, and the **occurs** attribute. The **name** attribute and the **type** attribute shall be specified. Permissible values and semantics of the **occurs** attribute are the same as in **ref** with the **label** attribute.

```
name  NMTOKEN  #REQUIRED
type  NMTOKEN  #REQUIRED
occurs CDATA #IMPLIED
```

The following content model describes the permissible content of **element**. In this content model, it is assumed that all element types representing facets are connected by "|" and declared as the value of the `facet` parameter entity.

```
(annotation?, (%facet;)*)
```

element is expanded to **ref**, **tag**, and **elementRule** according to the following rule.

a) A **ref** element shall be created, and the **element** shall be replaced by this **ref**. An appropriate label which does not collide with other labels shall be created as the value of the **label** attribute of this **ref**. The **occurs** attribute of the **element**, if any, shall be copied to the **ref**.

b) An **elementRule** element shall be created, and shall be added to this module. An appropriate role which does not collide with other roles shall be created as the value of the **role** attribute of the **elementRule**. The value of the **label** attribute shall be the label created in a). The hedge model of the **elementRule** shall be a datatype reference. The name of the referenced datatype shall be the value of the **type** attribute of the **element**. The content of the **elementRule** shall be the content of the **element**.

c) A **tag** element shall be created, and shall be added to this module. The value of the **role** attribute of this **tag** shall be the role created in b). The value of the **name** attribute of this **tag** shall be the value of the **name** attribute of the original **element**.

6.18 include

include provides a mechanism for referencing other modules.

include has the **moduleLocation** attribute.

```
moduleLocation  CDATA  #REQUIRED
```

The **moduleLocation** attribute references another module via a URI reference(IETF RFC 2396). The URI reference shall not contain a fragment identifier.

The following content model describes the permissible content of **include** elements:

```
(annotation?)
```

The module which contains **include** is said to be a *referencing module*, and a module referenced by the **moduleLocation** attribute is said to be a *referenced module.* The referenced module and the referencing module are required to specify the same value by the **targetNamespace** attribute.

6.19 div

div is introduced as a mechanism for grouping clauses, production rules, and **include** elements, and grouping of **export** elements in **interface** elements.

div has no attributes.

The following content model describes the permissible content of **div** elements. **div** may have subordinate **export** only when **interface** is an ancestor of the **div**. When a **div** element does not have an **interface** ancestor, it may contain subordinate clauses, production rules, and **include** elements.

```
(annotation?, div*,
  (((elementRule | hedgeRule | tag | attPool | include),
     (elementRule | hedgeRule | tag | attPool | include | div)*)
    |
    (export, (export | div)*))?)
```

6.20 annotation

annotation is introduced as a mechanism for embedding comments in modules.

annotation has no attributes.

The following content model describes the permissible content of **annotation** elements:

```
(appinfo | documentation)*
```

6.21 documentation

documentation is introduced as a note for human users to read.

documentation has the **source** attribute.

```
source  CDATA  #IMPLIED
```

When the **source** attribute is specified, the value shall be a URI reference (see IETF RFC 2396).

The following content model describes the permissible content of **documentation** elements:

```
(#PCDATA)
```

A **documentation** element specifying the **source** attribute shall be an empty element. Non-empty **documentation** elements shall not specify the **source** attribute.

6.22 appinfo

appinfo is a mechanism for embedding information for user programs that handle modules.

appinfo has the **source** attribute.

```
source  CDATA  #IMPLIED
```

When the **source** attribute is specified, the value shall be a URI reference (see IETF RFC 2396).

The following content model describes the permissible content of **appinfo** elements:

```
(#PCDATA)
```

An **appinfo** element specifying the **source** attribute shall be an empty element. Non-empty **appinfo** elements shall not specify the **source** attribute.

7 Datatypes

7.1 General

Datatypes in this Technical Report shall be either built-in datatypes of XML Schema Part 2 or datatypes specific to *RELAX Core*.

NOTE At present, users are not allowed to define new datatypes.

7.2 Built-in datatypes of XML Schema Part

7.2.1 string

This datatype represents the set of lexical representations of the **string** datatype of XML Schema Part 2. Applicable facets are the same as in the **string** datatype of XML Schema Part 2.

7.2.2 boolean

This datatype represents the set of lexical representations of the **boolean** datatype of XML Schema Part 2. Applicable facets are the same as in the **boolean** datatype of XML Schema Part 2.

7.2.3 float

This datatype represents the set of lexical representations of the **float** datatype of XML Schema Part 2. Applicable facets are the same as in the **float** datatype of XML Schema Part 2.

7.2.4 double

This datatype represents the set of lexical representations of the **double** datatype of XML Schema Part 2. Applicable facets are the same as in the **double** datatype of XML Schema Part 2.

7.2.5 decimal

This datatype represents the set of lexical representations of the **decimal** datatype of XML Schema Part 2. Applicable facets are the same as in the **decimal** datatype of XML Schema Part 2.

7.2.6 timeDuration

This datatype represents the set of lexical representations of the **timeDuration** datatype of XML Schema Part 2. Applicable facets are the same as in the **timeDuration** datatype of XML Schema Part 2.

7.2.7 recurringDuration

This datatype represents the set of lexical representations of the **recurringDuration** datatype of XML Schema Part 2. Applicable facets are the same as in the **recurringDuration** datatype of XML Schema Part 2.

7.2.8 binary

This datatype represents the set of lexical representations of the **binary** datatype of XML Schema Part 2. Applicable facets are the same as in the **binary** datatype of XML Schema Part 2.

7.2.9 uriReference

This datatype represents the set of lexical representations of the **uriReference** datatype of XML Schema Part 2. Applicable facets are the same as in the **uriReference** datatype of XML Schema Part 2.

7.2.10 ID

This datatype represents the set of lexical representations of the **ID** datatype of XML Schema Part 2. Applicable facets are the same as in the **ID** datatype of XML Schema Part 2.

This datatype can be referenced by the **type** attribute of **attribute**, but cannot be referenced by the **type** attribute of **elementRule** or **element**.

A **tag** element and those **attPool** elements referenced from this **tag** directly or indirectly shall not declare more than one attribute as **ID**.

If multiple **tag** elements sharing the same tag name declare **ID** attributes directly or indirectly, these attributes shall be declared by a single **attPool** element, which shall be referenced by each of these **tag** elements directly or indirectly.

In a document information item, more than one element information item shall not specify the same name as values of **ID** attributes. In other words, the value of an **ID** attribute shall uniquely identify an element information item.

7.2.11 IDREF

This datatype represents the set of lexical representations of the **IDREF** datatype of XML Schema Part 2. Applicable facets are the same as in the **IDREF** datatype of XML Schema Part 2.

This datatype can be referenced by the **type** attribute of **attribute**, but cannot be referenced by the **type** attribute of **elementRule** or **element**.

If multiple **tag** elements sharing the same tag name declare **IDREF** attributes directly or indirectly, these attributes shall be declared by **attPool** elements, each of which shall be referenced by each of these **tag** elements directly or indirectly.

The value of an **IDREF** attribute shall match the value of some **ID** attribute in the document information item.

7.2.12 ENTITY

This datatype represents the set of lexical representations of the **ENTITY** datatype of XML Schema Part 2. Applicable facets are the same as in the **ENTITY** datatype of XML Schema Part 2.

This datatype can be referenced by the **type** attribute of **attribute**, but cannot be referenced by the **type** attribute of **elementRule** or **element**.

The value of an **ENTITY** attribute shall match the name of some entity information item in the document information item.

7.2.13 NOTATION

This datatype represents the set of lexical representations of the **NOTATION** datatype of XML Schema Part 2. Applicable facets are the same as in the **NOTATION** datatype of XML Schema Part 2.

This datatype can be referenced by the **type** attribute of **attribute**, but cannot be referenced by the **type** attribute of **elementRule** or **element**.

The value of a **NOTATION** attribute shall match the name of some notation information item in the document information item.

7.2.14 QName

This datatype represents the set of lexical representations of the **QName** datatype of XML Schema Part 2. Applicable facets are the same as in the **QName** datatype of XML Schema Part 2.

7.2.15 language

This datatype represents the set of lexical representations of the **language** datatype of XML Schema Part 2. Applicable facets are the same as in the **language** datatype of XML Schema Part 2.

7.2.16 IDREFS

This datatype represents the set of lexical representations of the **IDREFS** datatype of XML Schema Part 2. Applicable facets are the same as in the **IDREFS** datatype of XML Schema Part 2.

This datatype can be referenced by the **type** attribute of **attribute**, but cannot be referenced by the **type** attribute of **elementRule** or **element**.

If multiple **tag** elements sharing the same tag name declare **IDREFS** attributes directly or indirectly, these attributes shall be declared by **attPool** elements, each of which shall be referenced by each of these **tag** elements directly or indirectly.

The value of an **IDREFS** attribute shall consist of names each of which shall match the value of some **ID** attribute in the document information item.

7.2.17 ENTITIES

This datatype represents the set of lexical representations of the **ENTITIES** datatype of XML Schema Part 2. Applicable facets are the same as in the **ENTITIES** datatype of XML Schema Part 2.

This datatype can be referenced by the **type** attribute of **attribute**, but cannot be referenced by the **type** attribute of **elementRule** or **element**.

The value of an **ENTITIES** attribute shall consist of names each of which shall match the name of some entity information item in the document information item.

7.2.18 NMTOKEN

This datatype represents the set of lexical representations of the **NMTOKEN** datatype of XML Schema Part 2. Applicable facets are the same as in the **NMTOKEN** datatype of XML Schema Part 2.

This datatype can be referenced by the **type** attribute of **attribute**, but cannot be referenced by the **type** attribute of **elementRule** or **element**.

7.2.19 NMTOKENS

This datatype represents the set of lexical representations of the **NMTOKENS** datatype of XML Schema Part 2. Applicable facets are the same as in the **NMTOKENS** datatype of XML Schema Part 2.

This datatype can be referenced by the **type** attribute of **attribute**, but cannot be referenced by the **type** attribute of **elementRule** or **element**.

7.2.20 Name

This datatype represents the set of lexical representations of the **Name** datatype of XML Schema Part 2. Applicable facets are the same as in the **Name** datatype of XML Schema Part 2.

7.2.21 NCName

This datatype represents the set of lexical representations of the **NCName** datatype of XML Schema Part 2. Applicable facets are the same as in the **NCName** datatype of XML Schema Part 2.

7.2.22 integer

This datatype represents the set of lexical representations of the **integer** datatype of XML Schema Part 2. Applicable facets are the same as in the **integer** datatype of XML Schema Part 2.

7.2.23 nonPositiveInteger

This datatype represents the set of lexical representations of the **nonPositiveInteger** datatype of XML Schema Part 2. Applicable facets are the same as in the **nonPositiveInteger** datatype of XML Schema Part 2.

7.2.24 negativeInteger

This datatype represents the set of lexical representations of the **negativeInteger** datatype of XML Schema Part 2. Applicable facets are the same as in the **negativeInteger** datatype of XML Schema Part 2.

7.2.25 long

This datatype represents the set of lexical representations of the **long** datatype of XML Schema Part 2. Applicable facets are the same as in the **long** datatype of XML Schema Part 2.

7.2.26 int

This datatype represents the set of lexical representations of the **int** datatype of XML Schema Part 2. Applicable facets are the same as in the **int** datatype of XML Schema Part 2.

7.2.27 short

This datatype represents the set of lexical representations of the **short** datatype of XML Schema Part 2. Applicable facets are the same as in the **short** datatype of XML Schema Part 2.

7.2.28 byte

This datatype represents the set of lexical representations of the **byte** datatype of XML Schema Part 2. Applicable facets are the same as in the **byte** datatype of XML Schema Part 2.

7.2.29 nonNegativeInteger

This datatype represents the set of lexical representations of the **nonNegativeInteger** datatype of XML Schema Part 2. Applicable facets are the same as in the **nonNegativeInteger** datatype of XML Schema Part 2.

7.2.30 unsignedLong

This datatype represents the set of lexical representations of the **unsingedLong** datatype of XML Schema Part 2. Applicable facets are the same as in the **unsingedLong** datatype of XML Schema Part 2.

7.2.31 unsignedInt

This datatype represents the set of lexical representations of the **unsignedInt** datatype of XML Schema Part 2. Applicable facets are the same as in the **unsignedInt** datatype of XML Schema Part 2.

7.2.32 unsignedShort

This datatype represents the set of lexical representations of the **unsignedShort** datatype of XML Schema Part 2. Applicable facets are the same as in the **unsignedShort** datatype of XML Schema Part 2.

7.2.33 unsingedByte

This datatype represents the set of lexical representations of the **unsignedByte** datatype of XML Schema Part 2. Applicable facets are the same as in the **unsignedByte** datatype of XML Schema Part 2.

7.2.34 positiveInteger

This datatype represents the set of lexical representations of the **positiveInteger** datatype of XML Schema Part 2. Applicable facets are the same as in the **positiveInteger** datatype of XML Schema Part 2.

7.2.35 timeInstant

This datatype represents the set of lexical representations of the **timeInstant** datatype of XML Schema Part 2. Applicable facets are the same as in the **timeInstant** datatype of XML Schema Part 2.

7.2.36 time

This datatype represents the set of lexical representations of the time datatype of XML Schema Part 2. Applicable facets are the same as in the time datatype of XML Schema Part 2.

7.2.37 timePeriod

This datatype represents the set of lexical representations of the **timePeriod** datatype of XML Schema Part 2. Applicable facets are the same as in the **timePeriod** datatype of XML Schema Part 2.

7.2.38 date

This datatype represents the set of lexical representations of the **date** datatype of XML Schema Part 2. Applicable facets are the same as in the **date** datatype of XML Schema Part 2.

7.2.39 month

This datatype represents the set of lexical representations of the **month** datatype of XML Schema Part 2. Applicable facets are the same as in the **month** datatype of XML Schema Part 2.

7.2.40 year

This datatype represents the set of lexical representations of the **year** datatype of XML Schema Part 2. Applicable facets are the same as in the **year** datatype of XML Schema Part 2.

7.2.41 century

This datatype represents the set of lexical representations of the **century** datatype of XML Schema Part 2. Applicable facets are the same as in the **century** datatype of XML Schema Part 2.

7.2.42 recurringDate

This datatype represents the set of lexical representations of the **recurringDate** datatype of XML Schema Part 2. Applicable facets are the same as in the **recurringDate** datatype of XML Schema Part 2.

7.2.43 recurringDay

This datatype represents the set of lexical representations of the **recurringDay** datatype of XML Schema Part 2. Applicable facets are the same as in the **recurringDay** datatype of XML Schema Part 2.

7.3 Datatypes Specific to RELAX

7.3.1 none

This datatype represents the empty set of strings. No character strings belong to this data type. This datatype has no applicable facets.

7.3.2 emptyString

This datatype represents a singleton set containing the empty string. This datatype has no applicable facets.

7.4 Facets

7.4.1 length

This facet is the same as the **length** facet of XML Schema Part 2.

7.4.2 minLength

This facet is the same as the **minLength** facet of XML Schema Part 2.

7.4.3 maxLength

This facet is the same as the **maxLength** facet of XML Schema Part 2.

7.4.4 pattern

This facet is the same as the **pattern** facet of XML Schema Part 2.

7.4.5 enumeration

This facet is the same as the **enumeration** facet of XML Schema Part 2.

7.4.6 maxInclusive

This facet is the same as the **maxInclusive** facet of XML Schema Part 2.

7.4.7 maxExclusive

This facet is the same as the **maxExclusive** facet of XML Schema Part 2.

7.4.8 minInclusive

This facet is the same as the **minInclusive** facet of XML Schema Part 2.

7.4.9 minExclusive

This facet is the same as the **minExclusive** facet of XML Schema Part 2.

7.4.10 presicion

This facet is the same as the **precision** facet of XML Schema Part 2.

7.4.11 scale

This facet is the same as the **scale** facet of XML Schema Part 2.

7.4.12 encoding

This facet is the same as the **encoding** facet of XML Schema Part 2.

7.4.13 duration

This facet is the same as the **duration** facet of XML Schema Part 2.

7.4.14 period

This facet is the same as the **period** facet of XML Schema Part 2.

8 Reference model

8.1 General

This Technical Report specifies the reference model in order to clarify the criteria for determining whether an island is compliant with a RELAX module and an optional element hedge model.

NOTE This reference model does not constrain implementations.

8.2 Creation of element hedge models

In the absence of an element hedge model for the top-level elements of an island, an element hedge model shall be created from the **label** attribute of the **export** elements in the given module.

First, for each label specified by the **label** attribute of these **export** elements, a **ref** element that reference to this label shall be created. Next, an element hedge model shall be created by wrapping these **ref** elements with a **choice** element. The **ref** elements and **choice** shall not have the **occurs** attribute.

8.3 Expansion of modules

Each **include** element shall be expanded to the referenced module. When a referenced module further references to another module, **include** elements in the referenced module shall be expanded in advance.

8.4 Expansion of element

element elements in element hedge models shall be expanded to **ref**, **elementRule** and **tag**.

8.5 Expansion of modules

Each **hedgeRef** element shall be replaced by element hedge models of the **hedgeRule** elements referenced by the **hedgeRef**. Details are as below:

a) Locate all **hedgeRule** elements that describe the label referenced by this **hedgeRef** element.

b) Wrap the hedge models of these **hedgeRule** elements by a **choice** element.

c) Copy the **occurs** attribute of the **hedgeRef** element to this **choice** element.

d) Replace the **hedgeRef** element with this **choice** element.

If other **hedgeRef** elements appear in the **choice** element, they shall be recursively expanded.

8.6 Expansion of tag embedded in elementRule

Each **tag** element embedded in **elementRule** shall be moved so that it becomes a sibling element of this **elementRule**. A role which does not collide with other roles shall be created, and shall be specified as the value of the **role** attribute of the **tag** and that of the **elementRule**. If the **tag** does not have the **name** attribute, it shall have the value of the **label** attribute of the **elementRule**.

8.7 Interpretation

An *interpretation* of a hedge is a mapping from each element in this hedge to a role and a label.

A hedge is compliant if it has at least one sound interpretation. An interpretation is sound if the following conditions hold:

a) Each element plays the associated role.

b) Derivation at each element is correct.

c) The sequence of labels associated with the top-level elements match the given element hedge model.

Consider an element e and its children $e_1, e_2, \ldots, e_n$. Let t_0 be the character sequence preceding e_1 in e, let t_i be the character sequence occurring between e_i and e_{i+1} in e, and let t_n be the character sequence following e_n in e. By definition, $t_0, e_1, t_1, e_2, t_2, \ldots, e_n, t_n$ provides the content of e.

Let l be the label associated with e, and let $l_1, l_2, \ldots, l_n$ be the labels associated with $e_1, e_2, \ldots, e_n$, respectively. Derivation at e is correct if there exists some **elementRule** such that the value of its label attribute is l, the value of its **role** attribute is the role associated with e, and $t_0, l_1, t_1, l_2, t_2, \ldots, l_n, t_n$ matches its hedge model.

At user option, the *RELAX Core* processor shall output message, when some of the attributes of e is not directly or indirectly (via some **attPool**) declared by the **tag** element describing the role associated with e.

Labels l_1 and l_2 not described by **hedgeRule** are said to be *contextually indistinguishable* when the following conditions hold:

a) The hedge model of some **elementRule** p_1 references to l_1.

b) The hedge model of some **elementRule** p_2 references to l_2.

c) Either p_1 and p_2 are identical, or their **role** and **label** attributes specify the same role and label, respectively.

NOTE 1 Every label is contextually indistinguishable from itself, if it is described by **elementRule** and referenced by some hedge model.

If it is possible to construct an element that plays both r_1 and r_2, they are said to be *coexistent*.

NOTE 2 Every role is coexistent with itself.

If the **role** and **label** attributes of some **elementRule** specify role r and label l, respectively, r is said to *lead to* l,.

The *RELAX Core* processor should but need not continue normal processing, when a RELAX module does not satisfy the following uniqueness condition.

a) If labels l_1 and l_2 are contextually indistinguishable, roles r_1 and r_2 are coexistent, and r_1 and r_2 lead to l_1 and l_2, respectively, then l_1 and l_2 are identical and r_1 and r_2 are identical.

The *RELAX Core* processor may output message and stop normal processing, when it receives a RELAX module not satisfying this uniqueness condition. The *RELAX Core* processor should output warning message when the uniqueness condition does not hold.

9 Conformance

9.1 General

This Technical Report defines conformance levels of RELAX modules and conformance levels of the *RELAX Core* processor.

9.2 Conformance levels of RELAX modules

This Technical Report defines two conformance levels of RELAX modules: "classic" and "fully relaxed".

The conformance level "classic" shall have restrictions as below:

a) **elementRule** shall not have the **label** attribute.

b) more than one **elementRule** shall not specify the same role.

c) more than one **hedgeRule** shall not specify the same label.

d) **tag** shall not specify the **role** attribute.

e) more than one **tag** shall not specify the same tag name.

f) **element** shall not exist

g) **tag** shall not exist as a child element of **elementRule** .

h) The child element of **mixed** shall be either **ref** with the **label** attribute, **hedgeRef**, or **choice**. They shall specify "*" as the value of the **occurs** attribute. When **choice** is the child element, item I) shall apply. When **hedgeRef** is the child element, item J) shall apply.

i) Child elements of those **choice** elements shown in item H) shall be either **ref** with the **label** attribute or **hedgeRef**. These child elements shall not specify the **occurs** attribute. When **hedgeRef** is a child element, item J) shall apply.

j) **hedgeRule** referenced by those **hedgeRef** elements shown in items H) and I) shall have either **ref** with the **label** attribute, **hedgeRef**, or **choice** as a hedge model. They shall not specify the **occurs** attribute. When **hedgeRef** is the hedge model, this item shall apply recursively.

k) Datatypes shall be restricted to **string**, **boolean**, **float**, **double**, **long**, **int**, **short**, **byte**, **ID**, **IDREF**, **ENTITY**, **NOTATION**, **IDREFS**, **ENTITIES**, **NMTOKEN**, and **NMTOKENS**.

l) Facets shall be restricted to **enumeration**, **maxInclusive**, **maxExclusive**, **minInclusive**, and **minExclusive**.

The conformance level "fully relaxed" shall not have any of the above restrictions.

9.3 Conformance levels of the *RELAX Core* processor

This Technical Report defines two conformance levels for the *RELAX Core* processor: "classic" and "fully relaxed".

The *RELAX Core* processor conforms to the conformance level "classic" if it can handle modules conforming to the conformance level "classic" correctly. Given a module containing features beyond the conformance level "classic", the *RELAX Core* processor may stop normal processing, and, at user option, provide appropriate message.

The *RELAX Core* processor conforms to the conformance level "fully relaxed" if it handles any RELAX module correctly.

When a RELAX module has syntactical errors (i.e., it is not a document information item or does not meet conditions specified in this Technical Report), further processing shall not occur. At user option, the *RELAX Core* processor shall report such syntactical errors.

Annex A

DTD for *RELAX Core*

A.1 The kernel of *RELAX Core*

```
<?xml version="1.0" encoding="utf-8"?>
<!--
DTD for RELAX Core (Ver 1.0)
$Id: RELAX Core.dtd 1.17 2000/09/11 13:24:35 murata Exp murata $
-->

<!ENTITY % datatype "User_defined_datatypes_are_disallowed_at_least_now">
<!ENTITY % datatype-definitions SYSTEM "datatypes.dtd">
%datatype-definitions;

<!--*******************************************************-->
<!--                                                       -->
<!--       The overall structure of RELAX modules.         -->
<!--                                                       -->
<!--*******************************************************-->

<!ELEMENT interface ((%annotation;)?, (export | div)*)>

<!ENTITY % clause "(tag|attPool)">

<!ENTITY % rule "(elementRule|hedgeRule)">

<!ELEMENT module ((%annotation;)?, interface?,
                   (%clause; | %rule; | div | include )*)>

<!ATTLIST module
              moduleVersion    CDATA     #IMPLIED
              RELAX CoreVersion CDATA     #REQUIRED
              targetNamespace  CDATA     #IMPLIED
              xmlns          CDATA     #FIXED "http://www.xml.gr.jp/xmlns/RELAX
Core"
>

<!--*******************************************************-->
<!--                                                       -->
<!--                    div                               -->
<!--                                                       -->
<!--*******************************************************-->

<!ELEMENT div ((%annotation;)?,
            div*,
            (((%rule; |%clause; | include ),
              (%rule; |%clause; | include | div)*)
             |
             (export, (export | div)*))?)>

<!--

(%rule; |%clause; | include | div)* is used when a div appears in a
module body, while (export | div)* is used when it appears in an
interface element.

 -->


<!--*******************************************************-->
<!--                                                      -->
<!--                  Interface                            -->
<!--                                                       -->
<!--*******************************************************-->
```

```
<!ELEMENT export ((%annotation;)?)>
<!ATTLIST export label NMTOKEN #REQUIRED>

<!--*******************************************************-->
<!--                                                       -->
<!--                   Include                             -->
<!--                                                       -->
<!--*******************************************************-->

<!ELEMENT include ((%annotation;)?)>
<!ATTLIST include moduleLocation CDATA #REQUIRED>


<!--*******************************************************-->
<!--                                                       -->
<!--                 Hedge Models                          -->
<!--                                                       -->
<!--*******************************************************-->

<!-- The parameter entity "particle" is used to describe element hedge
models.  It is also used as subordinates of <sequence>, <choice>,
and <mixed>. -->

<!ENTITY % particle "(ref | hedgeRef | choice | sequence | element
          | none | empty)">

<!ENTITY % hedgeModel
  "(%particle;  | mixed)">

<!-- The parameter entity "repeatable" is used to specify the "occurs"
 attribute, which is shared by several elements.  Permissible values
 are either "?", "+", or "*".  -->

<!ENTITY % repeatable '
              occurs      CDATA     #IMPLIED
'>


<!ELEMENT hedgeRef EMPTY >
<!ATTLIST hedgeRef
              label     NMTOKEN     #REQUIRED
              %repeatable;
>

<!ELEMENT ref EMPTY >
<!ATTLIST ref
              label     NMTOKEN     #IMPLIED
              role      NMTOKEN     #IMPLIED
              %repeatable;
>

<!ELEMENT empty EMPTY >

<!ELEMENT choice (%particle;)* >
<!ATTLIST choice
              %repeatable;
>

<!ELEMENT sequence (%particle;)* >
<!ATTLIST sequence
              %repeatable;
>

<!ELEMENT none EMPTY>

<!ELEMENT mixed (%particle;) >

<!ELEMENT element ((%annotation;)?, (%facet;)*)>
<!ATTLIST element
              name    NMTOKEN #REQUIRED
              type    NMTOKEN #REQUIRED
              %repeatable;
>


<!--*******************************************************-->
<!--                                                       -->
```

```
<!--                        Rules                          -->
<!--                                                       -->
<!--*******************************************************-->

<!ELEMENT elementRule ((%annotation;)?, tag?,
                   ((%hedgeModel;) | (%facet;)*))>

<!ATTLIST elementRule
     role      NMTOKEN #IMPLIED
     label     NMTOKEN #IMPLIED
     type      NMTOKEN #IMPLIED
>

<!ELEMENT hedgeRule  ((%annotation;)?, %particle;) >
<!ATTLIST hedgeRule
     label     NMTOKEN #REQUIRED
>


<!--*******************************************************-->
<!--                                                       -->
<!--                    Clauses                            -->
<!--                                                       -->
<!--*******************************************************-->

<!ENTITY % clauseBody "((%annotation;)?, ref*, attribute*)">

<!ELEMENT tag  (%clauseBody;)>
<!ATTLIST tag
              role    NMTOKEN     #IMPLIED
              name    NMTOKEN     #IMPLIED
>

<!ELEMENT attPool (%clauseBody;)>
<!ATTLIST attPool
              role  NMTOKEN     #REQUIRED
>

<!ELEMENT attribute ((%annotation;)?, (%facet;)*) >
<!ATTLIST attribute
              name     NMTOKEN     #REQUIRED
              required  (true)     #IMPLIED
              type     NMTOKEN     #IMPLIED
>
```

A.2 Datatypes

RELAX Core uses “datatypes.dtd” in Appendix B of XML Schema part 2, which defines the syntax for referencing datatypes and facets.

Annex B

RELAX Module for *RELAX Core*

B.1 Datatypes

```
<?xml version="1.0" encoding="utf-8"?>
<!DOCTYPE module SYSTEM "RELAX Core.dtd">
<!--
Module for RELAX Core (Ver 1.0)
$Id: RELAX Core.rlx 1.18 2000/09/11 13:49:57 murata Exp $

-->

<module
  moduleVersion="1.0"
  RELAX CoreVersion="1.0"
  targetNamespace="http://www.xml.gr.jp/xmlns/RELAX Core"
  xmlns="http://www.xml.gr.jp/xmlns/RELAX Core">

  <interface>
    <export label="module"/>
  </interface>

  <include moduleLocation="datatypes.rlx"/>

  <div>

    <annotation>
      <documentation>The overall structure of RELAX modules</documentation>
    </annotation>

    <elementRule role="module">
      <sequence>
  <ref label="annotation" occurs="?"/>
  <ref label="interface" occurs="?"/>
  <choice occurs="*">
    <hedgeRef label="clause"/>  <!-- forward references are fine -->
    <hedgeRef label="rule"/>    <!-- forward references are fine -->
    <ref label="divInModule"/>
    <ref label="include"/>
  </choice>
      </sequence>
    </elementRule>

    <tag name="module">
      <attribute name="moduleVersion" type="string"/>
      <attribute name="RELAX CoreVersion" type="string" required="true">
         <enumeration value="1.0"/>
      </attribute>
      <attribute name="targetNamespace" type="uriReference"/>
    </tag>

    <elementRule role="interface">
      <sequence>
        <ref label="annotation" occurs="?"/>
        <choice occurs="*">
          <ref label="export"/>
          <ref label="divInInterface"/>
        </choice>
      </sequence>
    </elementRule>

    <tag name="interface"/>

    <hedgeRule label="clause">
      <choice>
        <ref label="tag"/>
        <ref label="attPool"/>
      </choice>
    </hedgeRule>

    <hedgeRule label="rule">
      <choice>
        <ref label="elementRule"/>
        <ref label="hedgeRule"/>
      </choice>
    </hedgeRule>


    <elementRule label="divInModule">
      <annotation>
        <documentation>div elements in modules</documentation>
      </annotation>
      <tag name="div"/>
      <sequence>
        <ref label="annotation" occurs="?"/>
        <choice>
          <hedgeRef label="rule"/>
          <hedgeRef label="clause"/>
          <ref label="divInModule"/>
          <ref label="include"/>
        </choice>
      </sequence>
    </elementRule>
  </div>

  <div>
    <annotation>
      <documentation>Interface</documentation>
    </annotation>

    <elementRule role="export">
      <ref label="annotation" occurs="?"/>
    </elementRule>

    <tag name="export">
      <attribute name="label" required="true" type="NCName"/>
    </tag>

    <elementRule label="divInInterface">
      <annotation>
        <documentation>div elements in interfaces</documentation>
      </annotation>
      <tag name="div"/>
      <sequence>
        <ref label="annotation" occurs="?"/>
      <choice occurs="*">
          <ref label="export"/>
          <ref label="divInInterface"/>
        </choice>
      </sequence>
    </elementRule>

  </div>

  <div>
    <annotation>
      <documentation>Include</documentation>
    </annotation>

    <elementRule role="include">
      <ref label="annotation" occurs="?"/>
    </elementRule>

    <tag name="include">
      <attribute name="moduleLocation" type="uriReference" required="true"/>
    </tag>
  </div>

  <div>
    <annotation>
      <documentation>Hedge Models</documentation>
    </annotation>

    <hedgeRule label="particle">
      <annotation>
      <documentation>This is used to describe element hedge models.
It is also used as subordinates of sequence,
choice, and mixed.
</documentation>
    </annotation>
    <choice>
      <ref label="refWithLabel"/>
      <ref label="hedgeRef"/>
      <ref label="choice"/>
      <ref label="sequence"/>
      <ref label="element"/>
      <ref label="none"/>
      <ref label="empty"/>
    </choice>
  </hedgeRule>

  <hedgeRule label="hedgeModel">
    <choice>
      <hedgeRef label="particle"/>
      <ref label="mixed"/>
    </choice>
  </hedgeRule>

  <attPool role="repeatable">
    <annotation>
      <documentation>This is used to specify the "occurs" attribute,
which is shared by several elements.</documentation>
    </annotation>
    <attribute name="occurs" type="string">
      <enumeration value="?"/>
      <enumeration value="*"/>
      <enumeration value="+"/>
    </attribute>
  </attPool>

  <elementRule role="hedgeRef" type="emptyString"/>

  <tag name="hedgeRef">
    <ref role="repeatable"/>
    <attribute name="label" required="true" type="NCName"/>
  </tag>

  <elementRule label="refWithLabel" type="emptyString">
    <annotation>
      <documentation>ref elements with the label attribute</documentation>
    </annotation>
    <tag name="ref">
      <ref role="repeatable"/>
      <attribute name="label" required="true" type="NCName"/>
      <attribute name="role" type="none"/>
    </tag>
  </elementRule>

  <elementRule role="empty" type="emptyString"/>

  <tag name="empty"/>

  <elementRule role="choice">
    <hedgeRef label="particle" occurs="*"/>
  </elementRule>

  <tag name="choice">
    <ref role="repeatable"/>
  </tag>

  <elementRule role="sequence">
    <hedgeRef label="particle" occurs="*"/>
  </elementRule>

  <tag name="sequence">
    <ref role="repeatable"/>
  </tag>

  <elementRule role="none" type="emptyString"/>

  <tag name="none"/>
  <elementRule role="mixed">
    <hedgeRef label="particle"/>
  </elementRule>

  <tag name="mixed"/>

  <elementRule label="element">
    <annotation>
      <documentation>with the type attribute</documentation>
    </annotation>
    <tag name="element">
      <ref role="repeatable"/>
      <attribute name="name" required="true" type="NCName"/>
      <attribute name="type" required="true" type="NCName"/>
    </tag>
    <sequence>
      <ref label="annotation" occurs="?"/>
      <hedgeRef label="facet"  occurs="*"/>
    </sequence>
  </elementRule>
</div>

<div>

  <annotation>
    <documentation>Rules</documentation>
  </annotation>

  <elementRule label="elementRule">
    <annotation>
      <documentation>without an embedded tag element</documentation>
      <documentation>with the type attribute</documentation>
    </annotation>
    <tag name="elementRule">
      <attribute name="role" required="true" type="NCNAME"/>
      <attribute name="label" type="NCName"/>
      <attribute name="type" type="NCName" required="true"/>
    </tag>
    <sequence>
      <ref label="annotation" occurs="?"/>
      <hedgeRef label="facet" occurs="*"/>
    </sequence>
  </elementRule>

  <elementRule label="elementRule">
    <annotation>
      <documentation>with an embedded tag element</documentation>
      <documentation>with the type attribute</documentation>
    </annotation>
    <tag name="elementRule">
      <attribute name="role" type="none"/>
      <attribute name="label" required="true" type="NCName"/>
      <attribute name="type" type="NCName" required="true"/>
    </tag>
    <sequence>
      <ref label="annotation" occurs="?"/>
      <ref label="tagInRule"/>
      <hedgeRef label="facet" occurs="*"/>
    </sequence>
  </elementRule>

  <elementRule label="elementRule">
    <annotation>
      <documentation>without an embedded tag element</documentation>
      <documentation>with a hedge model</documentation>
    </annotation>
    <tag name="elementRule">
      <attribute name="role" required="true" type="NCName"/>
      <attribute name="label" type="NCName"/>
      <attribute name="type" type="none"/>
    </tag>
    <sequence>
      <ref label="annotation" occurs="?"/>
      <hedgeRef label="hedgeModel"/>
    </sequence>
  </elementRule>
  <elementRule label="elementRule">
    <annotation>
      <documentation>with an embedded tag element</documentation>
      <documentation>with a hedge model</documentation>
    </annotation>
    <tag name="elementRule">
      <attribute name="role" type="none"/>
      <attribute name="label" required="true" type="NCName"/>
      <attribute name="type" type="none"/>
    </tag>
    <sequence>
      <ref label="annotation" occurs="?"/>
      <ref label="tagInRule"/>
      <hedgeRef label="hedgeModel"/>
    </sequence>
  </elementRule>

  <elementRule role="hedgeRule">
    <sequence>
      <ref label="annotation" occurs="?"/>
      <hedgeRef label="hedgeModel"/>
    </sequence>
  </elementRule>

  <tag name="hedgeRule">
    <attribute name="label" required="true" type="NCName"/>
  </tag>

</div>

<div>

  <annotation>
    <documentation>Clauses</documentation>
  </annotation>

  <hedgeRule label="clauseBody">
    <sequence>
      <ref label="annotation" occurs="?"/>
      <ref label="refWithRole" occurs="*"/>
      <ref label="attribute" occurs="*"/>
    </sequence>
  </hedgeRule>

  <elementRule role="tag">
    <hedgeRef label="clauseBody"/>
  </elementRule>

  <tag name="tag">
    <attribute name="role" type="NCName"/>
    <attribute name="name" required="true" type="NCName"/>
  </tag>

  <elementRule label="tagInRule">
    <annotation>
      <documentation>tag elements embedded in   elementRules</documentation>
    </annotation>
    <tag name="tag">
      <attribute name="role" type="none"/>
      <attribute name="name" type="NCName"/>
    </tag>
    <hedgeRef label="clauseBody"/>
  </elementRule>

  <elementRule role="attPool">
    <hedgeRef label="clauseBody"/>
  </elementRule>

  <tag name="attPool">
    <attribute name="role" required="true" type="NCName"/>
  </tag>

  <elementRule label="refWithRole" type="emptyString">
    <annotation>
      <documentation>ref elements with the role attribute</documentation>
    </annotation>
    <tag name="ref">
```

```
        <attribute name="role" required="true" type="NCName"/>
        <attribute name="label" type="none"/>
      </tag>
    </elementRule>

    <elementRule role="attribute">
      <sequence>
        <ref label="annotation" occurs="?"/>
        <ref label="facet" occurs="*"/>
      </sequence>
    </elementRule>

    <tag name="attribute">
      <attribute name="name" required="true" type="NMTOKEN">
      <annotation>
          <documentation>Since RELAX Core is concerned with a single
    namespace, NCName might be more appropriate.  However,
    xml:space and xml:lang are so common that RELAX Core should
    allow them.
          </documentation>
        </annotation>
      </attribute>
      <attribute name="type" type="NCName"/>
      <attribute name="required" type="NMTOKEN">
        <enumeration value="true"/>
      </attribute>
    </tag>

  </div>

</module>
```

B.2 Datatypes

```
<?xml version="1.0" encoding="utf-8"?>
<!--
Module for XML Schemas: Part 2: Datatypes
$Id: datatypes.rlx 1.7 2000/09/13 17:13:46 murata Exp $

-->
<!DOCTYPE module SYSTEM "RELAX Core.dtd">
<module
       moduleVersion="1.0"
       RELAX CoreVersion="1.0"
  targetNamespace="http://www.xml.gr.jp/xmlns/RELAX Core"
       xmlns="http://www.xml.gr.jp/xmlns/RELAX Core">

  <hedgeRule label="minBound">
    <choice>
      <ref label="minInclusive"/>
      <ref label="minExclusive"/>
    </choice>
  </hedgeRule>

  <hedgeRule label="maxBound">
    <choice>
      <ref label="maxInclusive"/>
      <ref label="maxExclusive"/>
    </choice>
  </hedgeRule>

  <hedgeRule label="bounds">
    <choice>
      <hedgeRef label="minBound"/>
      <hedgeRef label="maxBound"/>
    </choice>
  </hedgeRule>

  <hedgeRule label="numeric">
    <choice>
      <ref label="precision"/>
      <ref label="scale"/>
    </choice>
  </hedgeRule>

  <hedgeRule label="ordered">
    <choice>
      <hedgeRef label="bounds"/>
      <hedgeRef label="numeric"/>
    </choice>
  </hedgeRule>

  <hedgeRule label="unordered">
    <choice>
      <ref label="pattern"/>
      <ref label="enumeration"/>
      <ref label="length"/>
      <ref label="maxLength"/>
      <ref label="minLength"/>
      <ref label="encoding"/>
      <ref label="period"/>
      <ref label="duration"/>
    </choice>
  </hedgeRule>

  <hedgeRule label="facet">
    <choice>
      <hedgeRef label="ordered"/>
      <hedgeRef label="unordered"/>
    </choice>
  </hedgeRule>

  <hedgeRule label="facetModel">
    <ref label="annotation" occurs="?"/>
  </hedgeRule>

  <attPool role="facetAttrs">
    <attribute name="value" required="true"/>
  </attPool>

<!--
RELAX Core published as JIS TR in 2000/5 does not have user-defined
types.  But this may change in the future.  Comments?

  <elementRule role="simpleType">
    <sequence>
      <ref label="annotation" occurs="?"/>
      <hedgeRef label="facet" occurs="*"/>
    </sequence>
  </elementRule>

  <tag name="simpleType">
    <attribute name="name" type="NMTOKEN"/>
    <attribute name="base" required="true" type="string"/>
  </tag>
-->

  <elementRule role="maxExclusive">
    <hedgeRef label="facetModel"/>
  </elementRule>

  <tag name="maxExclusive">
    <ref role="facetAttrs"/>
  </tag>

  <elementRule role="minExclusive">
    <hedgeRef label="facetModel"/>
  </elementRule>

  <tag name="minExclusive">
    <ref role="facetAttrs"/>
  </tag>

  <elementRule role="maxInclusive">
    <hedgeRef label="facetModel"/>
  </elementRule>

  <tag name="maxInclusive">
    <ref role="facetAttrs"/>
  </tag>
```

```
<elementRule role="minInclusive">
  <hedgeRef label="facetModel"/>
</elementRule>

<tag name="minInclusive">
  <ref role="facetAttrs"/>
</tag>

<elementRule role="precision">
  <hedgeRef label="facetModel"/>
</elementRule>

<tag name="precision">
  <attribute name="value" required="true" type="positive-integer"/>
</tag>

<elementRule role="scale">
  <hedgeRef label="facetModel"/>
</elementRule>

<tag name="scale">
  <attribute name="value" required="true" type="non-negative-integer"/>
</tag>

<elementRule role="length">
  <hedgeRef label="facetModel"/>
</elementRule>

<tag name="length">
  <attribute name="value" required="true" type="non-negative-integer"/>
</tag>

<elementRule role="minLength">
  <hedgeRef label="facetModel"/>
</elementRule>

<tag name="minLength">
  <attribute name="value" required="true" type="non-negative-integer"/>
</tag>

<elementRule role="maxLength">
  <hedgeRef label="facetModel"/>
</elementRule>

<tag name="maxLength">
  <attribute name="value" required="true" type="non-negative-integer"/>
</tag>

<elementRule role="enumeration">
  <hedgeRef label="facetModel"/>
</elementRule>

<tag name="enumeration">
  <ref role="facetAttrs"/>
</tag>

<elementRule role="pattern">
  <hedgeRef label="facetModel"/>
</elementRule>

<tag name="pattern">
  <ref role="facetAttrs"/>
</tag>

<elementRule role="encoding">
  <hedgeRef label="facetModel"/>
</elementRule>

<tag name="encoding">
  <attribute name="value" required="true" type="MNTOKEN">
    <enumeration value="hex">
<annotation>
  <documentation>each (8-bit) byte is encoded as a sequence
          of 2 hexidecimal digits</documentation>
</annotation>
    </enumeration>
```

```
      <enumeration value="base64">
  <annotation>
    <documentation>value  is  encoded  in  Base64  as  defined  in  the  MIME
RFC</documentation>
  </annotation>
      </enumeration>
    </attribute>
  </tag>

  <elementRule role="period">
    <hedgeRef label="facetModel"/>
  </elementRule>

  <tag name="period">
    <attribute name="value" required="true" type="timeDuration"/>
  </tag>

  <elementRule role="duration">
    <hedgeRef label="facetModel"/>
  </elementRule>

  <tag name="duration">
    <ref role="facetAttrs"/>
  </tag>

  <elementRule role="annotation">
    <choice occurs="*">
      <ref label="appinfo"/>
      <ref label="documentation"/>
    </choice>
  </elementRule>

  <tag name="annotation"/>

  <elementRule role="appinfo">
    <mixed><empty/></mixed>
  </elementRule>

  <tag name="appinfo">
    <attribute name="source" type="string"/>
  </tag>

  <elementRule role="documentation">
    <mixed><empty/></mixed>
  </elementRule>

  <tag name="documentation">
    <attribute name="source" type="string"/>
    <attribute name="xml:lang" type="string"/>
  </tag>

</module>
```

Bibliography

[1] W3C (World Wide Web Consortium), XML Schema Part 0: Premier, W3C Working Draft, http://www.w3.org/TR/xmlschema-0