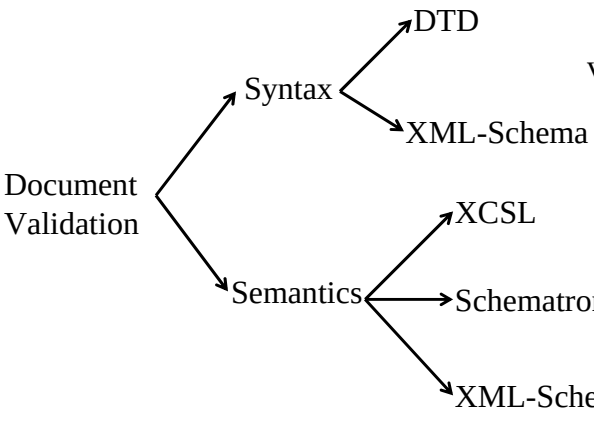
Grupo de Especificação e Processamento de Linguagens

Constraint Specification Languages:

comparing XCSL, Schematron and XML-Schemas

Constraint Specification Languages

Motivation



"Classical" validation of XML documents

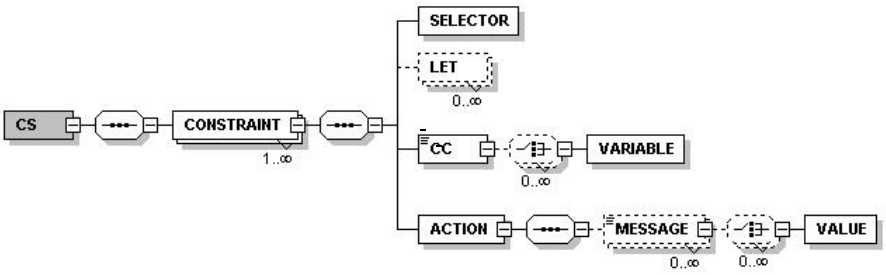
Constraint Specification Languages

Outline

- Brief description
  - XCSL
  - Schematron
  - XML-Schemas
- Semantic Constraints
  - Classification
  - Simplified Templates in XCSL/Schematron/XML-Schemas
- Case-studies
  - DTD
  - XML instance
  - Constraints in XCSL/Schematron/XML-Schemas
- Conclusion
- Future Work

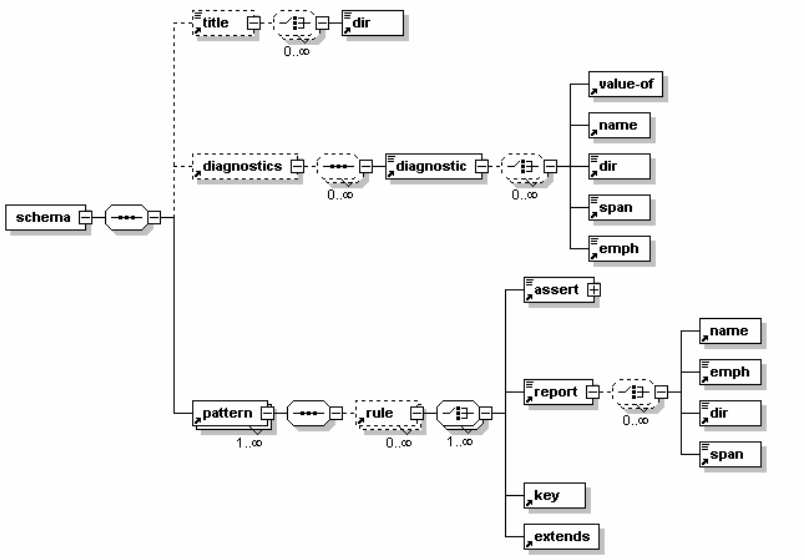
Constraint Specification Languages

XCSL



Constraint Specification Languages

Schematron



Constraint Specification Languages

XML-Schemas

Too complex to be shown here...

Constraint Specification Languages

Semantic Constraints

Constraint Specification Languages

Semantic Constraints (1)

- Case 1: Domain Range checking
- Case 2: Dependencies between two elements/attributes
- Case 3: Matching against a Regular Expression
- Case 4: Complex constraints

Constraint Specification Languages

Semantic Constraints (1a)

Colors used for the variable parts of the restrictions:

- Paths
- Element|Attribute names
- XPath Expressions
- Names of type of element
- Values
- Variables/lists of values of elements/attributes
- Messages
- Pattern titles

Constraint Specification Languages

Semantic Constraints (Case 1a)

- Domain range checking (XCSL restriction)

```
<CONSTRAINT>
   <SELECTOR SELEXP="path to the element"/>
   <CC>
     . | @attname < value
    </CC>
   <ACTION>
     <MESSAGE>
         Message...
         <VALUE SELEXP="path to any element/attribute |
         any expression applied to any element/attribute"/>
     </MESSAGE>
   </ACTION>
</CONSTRAINT>
```

Constraint Specification Languages

Semantic Constraints (Case 1b)

- Domain range checking (Schematron restriction)

```
<diagnostics>
   <diagnostic id="01">
       Message...
       <value-of select="path to any element/attribute |
       any expression applied to any element/attribute"/>
   </diagnostic>
</diagnostics>
<pattern name="pattern title">
   <rule context="path to the element">
      <assert test=" . | @attname < value"
      diagnostics="01"/>
   </rule>
</pattern>
```


Constraint Specification Languages

Semantic Constraints (Case 1c)

- Domain range checking (XML-Schema restriction)

```
<xs:simpleType name="name of type of the element/attribute">
      <xs:restriction base="xs:integer">
         <xs:minInclusive value="value"/>
      </xs:restriction>
</xs:simpleType>
```

Constraint Specification Languages

Semantic Constraints (Case 2a)

- Dependencies between two elements/attributes (XCSL restriction)

```
<CONSTRAINT>
   <SELECTOR SELEXP="path to the 1st element"/>
   <CC>
     . | @attname < path to the 2nd element/[ . | @attname]
    </CC>
   <ACTION>
     <MESSAGE>
          Message...
          <VALUE SELEXP="path to any element/attribute |
          any expression applied to any element/attribute"/>
     </MESSAGE>
   </ACTION>
</CONSTRAINT>
```


Constraint Specification Languages

Semantic Constraints (Case 2b)

- Dependencies between two elements/attributes (Schematron restriction)

```
<diagnostics>
   <diagnostic id="01">
       Message...
       <value-of select="path to any element/attribute |
       any expression applied to any element/attribute"/>
   </diagnostic>
</diagnostics>
<pattern name="pattern title">
   <rule context="path to the 1st element">
      <assert test=" . | @attname < path to the 2nd element/[ . | @attname] "
      diagnostics="01"/>
   </rule>
</pattern>
```

Constraint Specification Languages

Semantic Constraints (Case 2c)

- Dependencies between two elements/attributes (XML-Schema restriction)

Not specifiable...

Constraint Specification Languages

Semantic Constraints (Case 3a)

- Pattern Matching against a Regular Expression (XCSL restriction)

```
<CONSTRAINT>
   <SELECTOR SELEXP="path to the element"/>
   <CC>
      substring(.| @attname,i,n1)=literal_value and
      (string-length(number(substring(.| @attname,j,n2))) = value
   </CC>
   <ACTION>
      <MESSAGE>
         Message...
         <VALUE SELEXP="path to any element/attribute |
         any expression applied to any element/attribute"/>
      </MESSAGE>
   </ACTION>
</CONSTRAINT>
```

Values like:
Literal_value value_digits

Constraint Specification Languages

Semantic Constraints (Case 3b)

- Pattern Matching against a Regular Expression (Schematron restriction)

```
<diagnostics>
  <diagnostic id="01">
     Message...
     <value-of select="path to any element/attribute |
     any expression applied to any element/attribute"/>
  </diagnostic>
</diagnostics>
<pattern name="pattern title">
  <rule context="path to the element">
    <assert test="
          substring(.| @attname,i,n1)=literal_value and
          (string-length(number(substring(. | @attname,j,n2))) = value"
    diagnostics="01"/>
  </rule>
</pattern>
```

Values like:
Literal_value value_digits

Constraint Specification Languages

Semantic Constraints (Case 3c)

- Pattern Matching against a Regular Expression (XML-Schema restriction)

```
<xs:simpleType name="telement">
    <xs:restriction base="xs:string">
      <xs:pattern value="literal_value\d{value}"/>
    </xs:restriction>
  </xs:simpleType>
```

Values like:
Literal_value value_digits

Constraint Specification Languages

Semantic Constraints (Case 4a)

- Complex constraints –mixed content (XCSL restriction)

```
<CONSTRAINT>
   <SELECTOR SELEXP="path to the parent element"/>
   <CC>
      (count(elt1)=c_elt1) and (count(elt2)=c_elt2) and ... (count(eltn)=c_eltn)
      name(elo1[1]/following::*)='elo2' and
      name(elo2[1|2]/following::*)='elo3' and
         ...
   </CC>
   <ACTION>
      <MESSAGE>
         Message...
         <VALUE SELEXP="path to any element/attribute |
         any expression applied to any element/attribute"/>
      </MESSAGE>
   </ACTION>
</CONSTRAINT>
```

ATTENTION
elo2 may differ from *elt2* !!!

Constraint Specification Languages

Semantic Constraints (Case 4b)

- Complex constraints –mixed content (Schematron restriction)

```
<diagnostics>
   <diagnostic id="01">
      Message...
      <value-of select="path to any element/attribute |
      any expression applied to any element/attribute"/>
   </diagnostic>
</diagnostics>
<pattern name="pattern title">
   <rule context="path to the parent element">
      <assert test="
       (count(elt1)=c_elt1) and (count(elt2)=c_elt2) and ... (count(eltn)=c_eltn) a
       name(elo1[1]/following::*)='elo2' and
       name(elo2[1|2]/following::*)='elo3' and
          ..."
      diagnostics="01"/>
   </rule>
</pattern>
```

ATTENTION
elo2 may differ from *elt2* !!!

Constraint Specification Languages

Semantic Constraints (Case 4c)

- Complex constraints –mixed content (XML-Schema restriction)

```
<xs:complexType name="tparent element" mixed="true">
  <xs:sequence>
    <xs:element name="elo1" type="telo1" minOccurs="elo1Min" maxOccurs="elo1Max"/>
    <xs:element name="elo2" type="telo2" minOccurs="elo2Min" maxOccurs="elo2Max"/>
    ...
    <xs:element name="elon" type="telon" minOccurs="elonMin" maxOccurs="elonMax"/>
  </xs:sequence>
</xs:complexType>
```

ATTENTION
It may happen: *eloi*=*eloj* !!!

Constraint Specification Languages

Semantic Constraints (Case 4'a)

- Complex constraints – unicity problem (XCSL restriction)

```
<CONSTRAINT>
  <SELECTOR SELEXP="path to X branch"/>
  <LET NAME="nameKey1" VALUE="elementX | @attributeX"/>
  <CC>
    (count(path to Y branch[elementY | @attributeY = $nameKey1]) =
1)
  </CC>
  <ACTION>
   <MESSAGE>
      Message...
      element | @attribute:
      <VALUE SELEXP="$nameKey1"/>.
   </MESSAGE>
  </ACTION>
</CONSTRAINT>
```

Every value of *element | @attribute* that appears in the X branch exists in the Y branch.

Constraint Specification Languages

Semantic Constraints (Case 4'b)

- Complex constraints – unicity problem (Schematron restriction)

```
<diagnostics>
   <diagnostic id="01">
      Message...
      element | @attribute:
      <value-of select=" elementX | @attributeX "/>
</diagnostic>
</diagnostics>
<pattern abstract="true" id="uID">
   <rule context="path to Y branch">
      <key name="nameKey1" path="elementY | @attributeY"/>
   </rule>
</pattern>
<pattern name="pattern title">
   <rule context="path to X branch">
      <assert test="count(key('nameKey1',elementX | @attributeX ) = 1)"
      diagnostics="01"/>
</rule>
</pattern>
```

Every value of *element | @attribute* that appears in the X branch exists in the Y branch.

Constraint Specification Languages

Semantic Constraints (Case 4'c)

- Complex constraints – unicity problem (XML-Schema restriction)

Not specifiable...

Constraint Specification Languages

Semantic Constraints (5)

Kind of constraint \ Constraint Language	XCSL	Schematron	XML-Schemas
Domain Range checking	X	X	X
Dependencies between two elements/attributes	X	X	
Pattern Matching against a Regular Expression	X	X	X
Complex Constraints (mixed content)	X	X	X
Complex constraints (unicity problem)	X	X	

Easier

No personalized output

Constraint Specification Languages

Fiscal Certificate

Case-Studies

2nd Conference for a Divorce

Database

Conc

Constraint Specification Languages

Case-Study 1 – Fiscal Certificate (1)

- What is it?
- Problems it raises:
  - Dates
  - Department
  - Cardinality/order of mixed content elements' sub-elements

Constraint Specification Languages

Case-Study 1 – Fiscal Certificate (2)

- DTD:

```
<!ELEMENT fcert (header, body, ending)>
<!ELEMENT header (#PCDATA | department)*>
<!ELEMENT department (#PCDATA)>
<!ATTLIST department
   place CDATA "0101">
<!ELEMENT body (requester, request)>
<!ELEMENT requester (#PCDATA | name | CF | address)*>
<!ELEMENT name (#PCDATA)>
<!ELEMENT CF (#PCDATA)>
<!ELEMENT address (#PCDATA)>
<!ELEMENT request (#PCDATA | affinity | name | date | village |
parish | municipality)*>
<!ELEMENT affinity (#PCDATA)>
<!ELEMENT date (#PCDATA)>
<!ATTLIST date
   value CDATA "19000101">
<!ELEMENT village (#PCDATA)>
<!ELEMENT parish (#PCDATA)>
<!ATTLIST parish
   place CDATA "010101">
<!ELEMENT municipality (#PCDATA)>
<!ATTLIST municipality
   place CDATA "0101">
<!ELEMENT ending (#PCDATA | place | date)*>
<!ELEMENT place (#PCDATA)>
```

Constraint Specification Languages

Case-Study 1 – Fiscal Certificate (3)

- XML:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE fcert SYSTEM "fcert_cm.dtd">
<fcert>
   <header>
      Dear Sir, Chief of the Finance Department of
      <department place="110504">Lisbon's 4th Fiscal Parish</department>
   </header>
   <body>
      <requester>
         <name>Rita Santos </name>
         taxpayer Ner.
         <CF>31988455</CF>
         with the address
         <address>Pedras tortas Street, Ner 7 - 5423 Ranholas
        </address>
      </requester>
      <request>
         requests your Excellency to certify if, on behalf of the death of her
         ...
         <name>Francelestina Pereira e Santos</name>
         who died on the
         <date value="19990913">13th of September 1999</date>
         ...
```


Constraint Specification Languages

Case-Study 1 – Fiscal Certificate (3a)

- XML:

```
         parish of
         <parish place="100611">Salir de Matos</parish>
         municipality of
         <municipality place="1006">Caldas da Rainha</municipality>
         and maried she was with
         ...
      </request>
   </body>
   <ending>
      Ask that her request be granted
      <place>Caldas da Rainha</place>
      <date value="19991020">20th of October 1999</date>
      The requester
   </ending>
</fcert>
```

Constraint Specification Languages

Case-Study 1 – Fiscal Certificate (4)

- Problems it raises:
  - Dates
  - Department
  - Mixed Content

Constraint Specification Languages

Case-Study 1 – Fiscal Certificate (4a)

- XML:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE fcert SYSTEM "fcert_cm.dtd">
<fcert>
   ...
   <body>
      ...
      <request>
         requests your Excellency to certify if, on behalf of the death of her
         ...
         <name>Francelestina Pereira e Santos</name>
         who died on the
         <date value="20010803">3rd of August 2001</date>
         ...
      </request>
   </body>
   <ending>
      Ask that her request be granted
      <place>Caldas da Rainha</place>
      <date value="20010607">7th of June 2001</date>
      The requester
   </ending>
</fcert>
```

Constraint Specification Languages

Case-Study 1 – Fiscal Certificate (4b)

- XCSL restriction:

```
<CONSTRAINT>
   <SELECTOR SELEXP="//request/date"/>
   <CC>
      @value < /fcert/ending/date/@value
   </CC>
   <ACTION>
      <MESSAGE>
        The date of the death pointed out:
        <VALUE SELEXP="/fcert/body/request/date"/>, is
        posterior to the request date:
        <VALUE SELEXP="/fcert/ending/date"/>
      </MESSAGE>
   </ACTION>
</CONSTRAINT>
```

- Schematron restriction:

```
<diagnostics>
   <diagnostic id="00">
      Except!
   </diagnostic>
   <diagnostic id="01">
      The indicated date of the death:
      <value-of select="/fcert/body/request/date"/>,
      is posterior to the request date: <value-of select="/fcert/ending/date"/>
   </diagnostic>
</diagnostics>
<pattern name="dates">
   <rule context="//request/date">
      <assert test="@value < /fcert/ending/date/@value" diagnostics="01"/>
      <report test="@value < /fcert/ending/date/@value" diagnostics="00"/>
   </rule>
</pattern>
```


Constraint Specification Languages

Case-Study 1 – Fiscal Certificate (4c)

First attribute value - 20010803
Second one - 20010607

- XCSL error output:

```
<err-message>
        The date of the death pointed out: 3rd of August 2001,
        is posterior to the request date:  7th of June 2001
</err-message>
```

- Schematron error output:

```
dates
  The indicated date of the death: 3rd of August 2001 , is posterior to the  request date:
  7th of June de 2001
```

Constraint Specification Languages

Case-Study 1 – Fiscal Certificate (5)

- Problems it raises:
  - Dates
  - Department
  - Mixed Content

Constraint Specification Languages

Case-Study 1 – Fiscal Certificate (5a)

- XML:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE fcert SYSTEM "fcert_cm.dtd">
<fcert>
   <header>
      Dear Sir, Chief of the Finance Department of
      <department place="110504">Lisbon's 4th Fiscal Parish</department>
   </header>
   <body>
      ...
      <request>
         ...
         parish of
         <parish place="100611">Salir de Matos</parish>
         municipality of
         <municipality place="1006">Caldas da Rainha</municipality>
         and maried she was with
         ...
      </request>
   </body>
   ...
</fcert>
```

Case-Study 1 – Fiscal Certificate (5b)

Constraint Specification Languages

- XCSL restriction:

```
<CONSTRAINT>
  <SELECTOR SELEXP="//fcert/body/request"/>
  <CC>
     parish/@place = /fcert/header/department/@place
    or
    municipality/@place = /fcert/header/department/@place
  </CC>
  <ACTION>
    <MESSAGE>
      The request for this certificate shall not be delivered in this department
      Lisbon's 4th Fiscal Parish, but in the department in charge of the
      Salir de Matos's parish, Caldas da Rainha's municipality.
    </MESSAGE>
  </ACTION>
</CONSTRAINT>
```

- Schematron restriction:

```
<diagnostics>
  <diagnostic id="02">
    The request for this certificate shall not be delivered in this department
    Lisbon's 4th Fiscal Parish, but in the department in charge of the
    Salir de Matos's parish, Caldas da Rainha's municipality.
  </diagnostic>
</diagnostics>
<pattern name="Finance department">
  <rule context="//fcert/body/request">
    <assert test="parish/@place = /fcert/header/department/@place or
    municipality/@place = /fcert/header/department/@place" diagnostics="02"/>
  </rule>
</pattern>
```


Case-Study 1 – Fiscal Certificate (5c)

Constraint Specification Languages

```
department/@place=110504
parish/@place=100611
municipality/@place=1006
```

- XCSL error output:

```
<err-message>
    The request for this certificate shall not be delivered in this department
    Lisbon's 4th Fiscal Parish, but in the department in charge of the
    Salir de Matos's parish, Caldas da Rainha's municipality.
</err-message>
```

- Schematron error output:

```
Finance department
  The request for this certificate shall not be delivered in this department Lisbon's 4th
  Fiscal Parish, but in the department in charge of the Salir de Matos's parish,
  Caldas da Rainha 's municipality.
```

Constraint Specification Languages

Case-Study 1 – Fiscal Certificate (6)

- Problems it raises:
  - Dates
  - Department
  - Mixed Content (requester element)

Constraint Specification Languages

Case-Study 1 – Fiscal Certificate (6a)

- XML:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE fcert SYSTEM "fcert_cm.dtd">
<fcert>
   ...
   <body>
      <requester>
         <name>Rita Santos </name>
         taxpayer Ner.
         <CF>31988455</CF>
         with the address
         <address>Pedras tortas Street, Ner 7 - 5423 Ranholas
        </address>
      </requester>
      ...
   </body>
</fcert>
```

Case-Study 1 – Fiscal Certificate (6b)

Constraint Specification Languages

- XCSL restriction:

```
<CONSTRAINT>
  <SELECTOR SELEXP="//certibody/requester"/>
  <CC>
    (count(name) = 1) and
    (count(CF) = 1) and
    (count(address) = 1) and
    name(name[1]/following::*) ='CF' and
    name(CF[1]/following::*) ='address'
  </CC>
  <ACTION>
    <MESSAGE>
      Either -requester- sub-elements occur in a wrong order,
      either they occur a wrong number of times.
    </MESSAGE>
  </ACTION>
</CONSTRAINT>
```

- XML-Schema restriction:

```
<xs:complexType name="trequester" mixed="true">
  <xs:sequence>
    <xs:element name="name" type="tname"/>
    <xs:element name="CF" type="tCF"/>
    <xs:element name="address" type="taddress"/>
  </xs:sequence>
</xs:complexType>
```

- Schematron restriction:

```
<diagnostics>
  <diagnostic id="04"
    Either -requester- sub-elements occur in a wrong order, either
    they occur a wrong number of times.
  </diagnostic>
</diagnostics>
<pattern name="requester element">
  <rule context="//certi/body/requester">
    <assert test="(count(name) = 1) and
      (count(CF) = 1) and
      (count(address) = 1) and
      name(name[1]/following::*) ='CF' and
      name(CF[1]/following::*) ='address'"
      diagnostics="04"/>
  </rule>
</pattern>
```


Case-Study 1 – Fiscal Certificate (6c)

Constraint Specification Languages

If the XML instance had two name elements

- XCSL error output:

```
<err-message>
    Either -requester- sub-elements occur in a wrong order,
  either  they occur a wrong number of times.
</err-message>
```

- Schematron error output:

```
requester element
    Either -requester- sub-elements occur in a wrong order,
either  they occur a wrong number of times.
```

- XML-Schema error output:

```
document  invalid...
```

Menu

Case-Study 2 – 2nd Conference for a Divorce (1)

- What is it?
- Problems it raises:
  - Days since the first petition

Case-Study 2 – 2nd Conference for a Divorce (2)

- DTD:

```
<!ELEMENT div_2c (header, body, ending)>
<!ELEMENT header (sender, addressee)>
<!ELEMENT sender (#PCDATA | cdepart)*>
<!ELEMENT cdepart (#PCDATA)>
<!ELEMENT addressee (#PCDATA | court)*>
<!ELEMENT court (#PCDATA)>
<!ELEMENT body (requesters, request)>
<!ELEMENT requesters (#PCDATA | name)*>
<!ELEMENT name (#PCDATA)>
<!ELEMENT request (#PCDATA | date | article)*>
<!ELEMENT date (#PCDATA)>
<!ATTLIST date
   value CDATA "19000101"    >
<!ELEMENT article (#PCDATA)>
<!ELEMENT ending (text, place, date, signature, signature)>
<!ELEMENT place (#PCDATA)>
<!ELEMENT signature (#PCDATA)>
<!ELEMENT text (#PCDATA)>
```

Constraint Specification Languages

Case-Study 2 – 2nd Conference for a Divorce (3)

- XML:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE div_2c SYSTEM "div_2c02.dtd">
<div_2c>
 <header>
  ...
 </header>
 <body>
  ...
  <request>
   identified in the referred Action of Divorce official papers,
   having accomplished the first conference in the
   <date value="20010406">6th of April of 2001</date>
   and both maintaining their will to divorce, come,
   ...
  </request>
 </body>
 <ending>
   <date value="20010506">6th of May of 2001</date>
  ...
 </ending>
</div_2c>
```


Constraint Specification Languages

Case-Study 2 – 2nd Conference for a Divorce (4)

- Problems it raises:
  - Days since the first petition

Case-Study 2 – 2nd Conference for a Divorce (4a)

Constraint Specification Languages

- XML:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE div_2c SYSTEM "div_2c02.dtd">
<div_2c>
 <header>
  ...
 </header>
 <body>
  ...
  <request>
   identified in the referred Action of Divorce official papers,
   having accomplished the first conference in the
   <date value="20010406">6th of April of 2001</date>
   and both maintaining their will to divorce, come,
   ...
  </request>
 </body>
 <ending>
   <date value="20010506">6th of May of 2001</date>
  ...
 </ending>
</div_2c>
```


Case-Study 2 – 2nd Conference for a Divorce (4b)

Constraint Specification Languages

- XCSL restriction:
- Schematron restriction:

```
<diagnostics>
 <diagnostic id="01">
  Less than 90 days undergone since the first conference.
  You will have to wait a little longer!!
 </diagnostic>
</diagnostics>
<pattern name="Days since the First Conference">
 <rule context="//div_2c">
  <assert test="(((substring(ending/date/@value,7,2)+ floor((153*(substring(ending/date/@value,5,2)+
   12 * (floor((14-(substring(ending/date/@value,5,2))) div 12))-3)+2) div 5)+
   (365 * (substring(ending/date/@value,1,4)+4800-
    (floor((14-(substring(ending/date/@value,5,2))) div 12)))))+floor((substring(ending/date/@value,1,4)+
   4800-((floor((14-(substring(ending/date/@value,5,2))) div 12)) div 4)-
   floor((substring(ending/date/@value,1,4)+4800- (floor((14-(substring(ending/date/@value,5,2)))
    div 12)) div 100)+
   floor((substring(ending/date/@value,1,4)+4800- (floor((14-(substring(ending/date/@value,5,2)))
    div 12)) div 400) -32045)
   - ...)
   >= 90" diagnostics="01">
  </assert>
 </rule>
</pattern>
```

```
<CONSTRAINT>
 <SELECTOR SELEXP="//div_2c"/>
 <LET NAME="a" VALUE="(floor((14-substring(ending/date/@value,5,2)) div 12))"/>
 <LET NAME="y" VALUE="(substring(ending/date/@value,1,4) + 4800 - $a)"/>
 <LET NAME="m" VALUE="(substring(ending/date/@value,5,2) + 12 * $a - 3)"/>
 <LET NAME="t" VALUE="(substring(ending/date/@value,7,2)
                       + floor((153 * $m + 2) div 5) +
                       (365 * $y) + floor($y div 4) -
                       floor($y div 100) +
                       floor($y div 400) - 32045)"/>
 ...
 <CC>
  ($t - $t2) >= 90
 </CC>
 <ACTION>
  <MESSAGE>
   Only <VALUE SELEXP="($t - $t2)"/> days undergone since the first conference...
   You will have to wait a little longer!!
  </MESSAGE>
 </ACTION>
</CONSTRAINT>
```

Constraint Specification Languages

Case-Study 2 – 2nd Conference for a Divorce (4c)

First attribute value - 20010406
Second one - 200105606

- XCSL error output:

```
<err-message>
   Only 30 days undergone
   since the first conference...
   You will have to wait a little longer!!
</err-message>
```

- Schematron error output:

```
Days since the First Conference
  Less than 90 days undergone since the first conference... You will have
  to wait a little longer!!
```

Menu

Constraint Specification Languages

Case-Study 3 – Database (1)

- What is it?
- Problems it raises:
  - key field
  - every record defined in the STRUCTURE sub-tree is used to instantiate the records in the DATA sub-tree
  - every record in the DATA sub-tree uses fields defined in the STRUCTURE sub-tree

Constraint Specification Languages

Case-Study 3 – Database (2)

• DTD:

```
<!ELEMENT DB (STRUCTURE, DATA)>
<!ELEMENT STRUCTURE (TABLE)+>
<!ELEMENT TABLE (COLUMNS, KEYS)>
<!ATTLIST TABLE
   NAME CDATA #REQUIRED>
<!ELEMENT COLUMNS (COLUMN)+>
<!ELEMENT COLUMN EMPTY>
<!ATTLIST COLUMN ...>
<!ELEMENT KEYS (PKEYS)>
<!ELEMENT PKEYS (PKEY)+>
<!ATTLIST PKEYS TYPE (simple | complex) #REQUIRED>
<!ELEMENT PKEY EMPTY>
<!ATTLIST PKEY NAME CDATA #REQUIRED>
<!ELEMENT DATA (items)+>
<!ELEMENT items (items-REG+)>
<!ATTLIST items NAME CDATA #REQUIRED>
<!ELEMENT items-REG (FIELD)+>
<!ELEMENT FIELD (#PCDATA)>
<!ATTLIST FIELD name CDATA #REQUIRED>
```


Constraint Specification Languages

Case-Study 3 – Database (3)

• XML:

```
<?xml version="1.0"?>
<!DOCTYPE DB SYSTEM "dbml_g.dtd">
<DB>
   <STRUCTURE>
      <TABLE NAME="stocks"> ... </TABLE>
      <TABLE NAME="suppliers"> ... </TABLE>
      <TABLE NAME="clients">
         <COLUMNS>
            <COLUMN NAME="cclient" TYPE="nvarchar" SIZE="10" NULL="no"/>
            <COLUMN NAME="name" TYPE="nvarchar" SIZE="50" NULL="no"/>
            <COLUMN NAME="contact" TYPE="nvarchar" SIZE="10" NULL="no"/>
            <COLUMN NAME="account" TYPE="nvarchar" SIZE="10" NULL="no"/>
         </COLUMNS>
         <KEYS>
            <PKEYS TYPE="simple">
               <PKEY NAME="cclient"/>
            </PKEYS>
         </KEYS>
      </TABLE>
      <TABLE NAME="orders"> ... </TABLE>
   </STRUCTURE>
```

Constraint Specification Languages

Case-Study 3 – Database (3a)

• XML:

```
<DATA>
  <items NAME="stocks"> ... </items>
  <items NAME="suppliers"> ... </items>
  <items NAME="clients">
    <items-REG>
      <FIELD name="cclient">c001</FIELD>
      <FIELD name="name">Corner's Cafe</FIELD>
      <FIELD name="contact">123456324</FIELD>
      <FIELD name="account">1234567890123456789O1</FIELD>
    </items-REG>
    <items-REG>
      <FIELD name="cclient">c002</FIELD>
      <FIELD name="name">Supermimo Supermarket</FIELD>
      <FIELD name="account">09876543210987654321O</FIELD>
    </items-REG>
    ...
  </items>
  <items NAME="orders"> ... </items>
</DATA>
</DB>
```


Constraint Specification Languages

Case-Study 3 – Database (4)

• Problems it raises:

- key field
- every field defined in the STRUCTURE sub-tree is used to instantiate the records in the DATA sub-tree
- every record in the DATA sub-tree uses fields defined in the STRUCTURE sub-tree

Case-Study 3 – Database (4a)

Constraint Specification Languages

• XML:

```
<?xml version="1.0"?>
<!DOCTYPE DB SYSTEM "dbml_g.dtd">
<DB>
  <STRUCTURE> ...
    <TABLE NAME="clients"> <COLUMNS>
        <COLUMN NAME="cclient" TYPE="nvarchar" SIZE="10" NULL="no"/>
        <COLUMN NAME="name" TYPE="nvarchar" SIZE="50" NULL="no"/>
        <COLUMN NAME="contact" TYPE="nvarchar" SIZE="10" NULL="no"/>
        <COLUMN NAME="account" TYPE="nvarchar" SIZE="10" NULL="no"/>
      </COLUMNS> <KEYS> <PKEYS TYPE="simple">
                          <PKEY NAME="cclient"/>
                        </PKEYS>
      </KEYS> </TABLE>
    ...
  </STRUCTURE>
  <DATA> ...
    <items NAME="clients">
      <items-REG>
        <FIELD name="cclient">c002</FIELD>
        <FIELD name="name">Supermimo Supermarket</FIELD>
        <FIELD name="account">09876543210987654321O</FIELD>
      </items-REG>
      ... </items> ...
  </DATA>
</DB>
```


Case-Study 3 – Database (4b)

Constraint Specification Languages

• XCSL restriction:

```
<CONSTRAINT>
  <SELECTOR SELEXP="TABLE[@NAME='clients']/COLUMNS/COLUMN"/>
  <LET NAME="tableclients" VALUE="@NAME"/>
  <CC>
    (count(//items[@NAME='clients']/items-REG/FIELD[@name = $tableclients]) =
     count(//items[@NAME='clients']/items-REG))
  </CC>
  <ACTION>
    <MESSAGE>WARNING:
      The field <VALUE SELEXP="$tableclients"/> was not used in every record of the "clients" table
      (or was used more than once in some record).
    </MESSAGE>
  </ACTION>
</CONSTRAINT>
```

• Schematron restriction:

```
<diagnostics>
  <diagnostic id="03a">
    The field <value-of select="@NAME"/> was not used in every
    record of the "clients" table (or was used more than once in some record).
  </diagnostic>
</diagnostics>
<pattern abstract="true" id="ch">
  <rule context="items[@NAME='clients']/items-REG/FIELD">
    <key name="tableclients" path="@name"/>
  </rule>
</pattern>
<pattern name="Clients table (use of all the defined fields)">
  <rule context="TABLE[@NAME='clients']/COLUMNS/COLUMN">
    <assert test="(count(key('tableclients',@NAME))) =
    count(//items[@NAME='clients']/items-REG)" diagnostics="03a">
    </assert>
  </rule>
</pattern>
```

Case-Study 3 – Database (4c)

STRUCTURE sub-tree – cclient, name, contact, account
DATA sub-tree – cclient, name, account

- XCSL error output:

```
<err-message>WARNING:
      The field contact was not used in every record of the
      "clients" table (or was used more than once in some record).
 </err-message>
```

- Schematron error output:

```
Clients table (use of all the defined fields)
 /DB[1]/STRUCTURE[1]/TABLE[3]/COLUMNS[1]/COLUMN[3]
 <COLUMN NAME="contact" TYPE="nvarchar" SIZE="10" NULL="no">...</>
 The field contact was not used in every record of the "clients" table
  (or was used more than once in some record).
```

Menu

Conclusion

- *Do they do the same job?*
- *Are there some kind of constraints that are easier to specify with one of them?*
- *Do you need different background to use the tools?*
- *Is it possible to use them in similar situations (the same DTD, the same XML instances)?*
- *May we use them to produce an equal result?*
- *How do XCSL and Schematron relate to XML Schemas?*
- *What is the intersection area of these three?*
- *What kind of constraints each one of these three is able to specify?*

Constraint Specification Languages

Future Work

- Third Generation Stylesheets:
  abstracting from constraint templates
- Comparison with Xpath 2.0:
  exploring new trends

Grupo de Especificação e Processamento de Linguagens

The End

Marta Jacinto marta.jacinto@itij.mj.pt
Giovani Librelotto grl@di.uminho.pt
José Carlos Ramalho jcr@di.uminho.pt
Pedro Rangel Henriques prh@di.uminho.pt