

03 Document Type Definition (DTD)

Document Type Definition (DTD)

© Janusz R. Getta CSCI235/MCS9235/CSCI835 Databases, SCIT, Autumn 2015 1

03 Document Type Definition (DTD)

Document Type Definition

Document Type Definition (DTD) defines the structures of XML documents

Document Type Definition (DTD) uses the notation of **regular expressions** (see Appendix 1) to define XML documents

Document Type Definition (DTD) consists of the following components: **name of root element**, **element declarations**, **attribute lists**, **entity declarations**

© Janusz R. Getta CSCI235/MCS9235/CSCI835 Databases, SCIT, Autumn 2015 2

03 Document Type Definition (DTD)

Document Type Definition

Components of DTD:

Name of root element and document type declarations

```
<!DOCTYPE root-element [ doctype-declaration... ]>
```

Element declaration

```
<!ELEMENT element-name content-model>
```

Attribute list

```
<!ATTLIST element-name attr-name attr-type attr-default ...>
```

Entity declaration

```
<!ENTITY entity-name "entity-value">
```

© Janusz R. Getta CSCI235/MCS9235/CSCI835 Databases, SCIT, Autumn 2015 3

03 Document Type Definition (DTD)

Name of root element and document type declaration

Internal DTD declaration

```
<!DOCTYPE root-element [ doctype-declaration... ]>
```

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE book [
  <!ELEMENT book (title, chapter+)>
  <!ATTLIST book author CDATA #REQUIRED>
  <!ELEMENT title (#PCDATA)>
  <!ELEMENT chapter (#PCDATA)>
  <!ATTLIST chapter id ID #REQUIRED>
]>
<book author="C.J.Date">
  <title>Databases</title>
  <chapter id="I">Introduction</chapter>
  <chapter id="II">Conceptual modeling</chapter>
  <chapter id="III">SQL</chapter>
</book>
```

Name of root element
bookStructure of root
element bookAttribute of root
element book

Structure element title

© Janusz R. Getta CSCI235/MCS9235/CSCI835 Databases, SCIT, Autumn 2015 4

03 Document Type Definition (DTD)

Name of root element and document type declaration

External DTD declaration

```
<!DOCTYPE root-element SYSTEM "file.dtd">
```

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE book SYSTEM "book.dtd">
<book author="C.J.Date">
  <title>Databases</title>
  <chapter id="I">Introduction</chapter>
  <chapter id="II">Conceptual modeling</chapter>
  <chapter id="III">SQL</chapter>
</book>
```

Location of DTD

book.dtd

```
<!ELEMENT book (title, chapter+)>
<!ATTLIST book author CDATA #REQUIRED>
<!ELEMENT title (#PCDATA)>
<!ELEMENT chapter (#PCDATA)>
<!ATTLIST chapter id ID #REQUIRED>
```

© Janusz R. Getta CSCI235/MCS9235/CSCI835 Databases, SCIT, Autumn 2015 5

03 Document Type Definition (DTD)

Element declaration

```
<!ELEMENT element-name content-model>
```

Content models

EMPTY	:no contents is allowed
ANY	:any contents is allowed
(#PCDATA element-name ...)*	: "mixed content", arbitrary sequence of character data and listed elements
regular expression over the element names	: sequence of elements matching the expression
choice	: (...)
sequence	: (... , ... , ...)
optional	: ...?
zero or more	: ...*
one or more	: ...+

#PCDATA means "Parsed Character Data", i.e. text in the XML document following the element tag parsed when looking for more markup tags

© Janusz R. Getta CSCI235/MCS9235/CSCI835 Databases, SCIT, Autumn 2015 6

03 Document Type Definition (DTD)

Example of element declaration

Content model EMPTY

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE book [
<!ELEMENT book (title,void)>
<!ELEMENT title (#PCDATA)>
<!ELEMENT void EMPTY>
]>
<book>
  <title>Databases</title>
  <void/>
</book>
```

© Janusz R. Getta CSCI235/MCS9235/CSCI835 Databases, SCIT, Autumn 2015 7

03 Document Type Definition (DTD)

Example of element declaration

Content model ANY

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE book [
<!ELEMENT book (title,whatever)>
<!ELEMENT title (#PCDATA)>
<!ELEMENT whatever ANY>
]>
<book>
  <title>Databases</title>
  <whatever>125347 28 72$#^$^asd%$%jj+ @asgdg</whatever>
</book>
```

© Janusz R. Getta CSCI235/MCS9235/CSCI835 Databases, SCIT, Autumn 2015 8

03 Document Type Definition (DTD)

Example of element declaration

Content model "mixed"

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE book [
<!ELEMENT book (title,comment)>
<!ELEMENT title (#PCDATA)>
<!ELEMENT reference (#PCDATA)>
<!ELEMENT comment (#PCDATA|reference)*>
]>
<book>
  <title>Databases</title>
  <comment>This book frequently refers to
    <reference>Ullman200</reference>
    in chapter one.</comment>
</book>
```

© Janusz R. Getta CSCI235/MCS9235/CSCI835 Databases, SCIT, Autumn 2015 9

03 Document Type Definition (DTD)

Example of element declaration

Content model "regular expression"

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE book [
<!ELEMENT book (title,(subtitle,(section)*)+)>
<!ELEMENT title (#PCDATA)>
<!ELEMENT subtitle (#PCDATA)>
<!ELEMENT section (#PCDATA)>
]>
<book>
  <title>Databases</title>
  <subtitle>Introduction</subtitle>
  <section>Section 1.1</section>
  <section>Section 1.2</section>
  <subtitle>Data modeling</subtitle>
  <section>Section 2.1</section>
  <subtitle>SQL</subtitle>
</book>
```

© Janusz R. Getta CSCI235/MCS9235/CSCI835 Databases, SCIT, Autumn 2015 10

03 Document Type Definition (DTD)

Example of element declaration

Content model "choice"

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE students [
<!ELEMENT students (student*)>
<!ELEMENT student (snum,(postgraduate|undergraduate))>
<!ELEMENT snum (#PCDATA)>
<!ELEMENT undergraduate (#PCDATA)>
<!ELEMENT postgraduate (#PCDATA)>
]>
<students>
  <student>
    <snum>324777</snum>
    <undergraduate>Computer Science</undergraduate>
  </student>
  <student>
    <snum>235778</snum>
    <postgraduate>Computer Science</postgraduate>
  </student>
</students>
```

© Janusz R. Getta CSCI235/MCS9235/CSCI835 Databases, SCIT, Autumn 2015 11

03 Document Type Definition (DTD)

Example of element declaration

Content model "sequence"

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE student [
<!ELEMENT student (snum,first-name,initials,last-name,country)>
<!ELEMENT snum (#PCDATA)>
<!ELEMENT first-name (#PCDATA)>
<!ELEMENT initials (#PCDATA)>
<!ELEMENT last-name (#PCDATA)>
<!ELEMENT country (#PCDATA)>
]>
<student>
  <snum>1234567</snum>
  <first-name>Janusz</first-name>
  <initials>R</initials>
  <last-name>Getta</last-name>
  <country>Oz</country>
</student>
```

© Janusz R. Getta CSCI235/MCS9235/CSCI835 Databases, SCIT, Autumn 2015 12

03 Document Type Definition (DTD)

Example of element declaration

Content model "optional"

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE person [
  <!ELEMENT person (first-name,initials?,last-name,
    date-of-birth?,phone?,email?)>
  <!ELEMENT first-name (#PCDATA)>
  <!ELEMENT initials (#PCDATA)>
  <!ELEMENT last-name (#PCDATA)>
  <!ELEMENT date-of-birth (#PCDATA)>
  <!ELEMENT phone (#PCDATA)>
  <!ELEMENT email (#PCDATA)>
]>
<person>
  <first-name>Janusz</first-name>
  <last-name>Getta</last-name>
  <email>jrg@uow.edu.au</email>
</person>
```

© Janusz R. Getta CSC1235/MCS9235/CSC1835 Databases, SCIT, Autumn 2015

13

03 Document Type Definition (DTD)

Example of element declaration

Content model "zero or more"

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE name [
  <!ELEMENT name (letter,(letter|digit)*)>
  <!ELEMENT letter (#PCDATA)>
  <!ELEMENT digit (#PCDATA)>
]>
<name>
  <letter>J</letter>
</name>
```

© Janusz R. Getta CSC1235/MCS9235/CSC1835 Databases, SCIT, Autumn 2015

14

03 Document Type Definition (DTD)

Example of element declaration

Content model "one or more"

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE name [
  <!ELEMENT name (letter,(letter|digit)+)>
  <!ELEMENT letter (#PCDATA)>
  <!ELEMENT digit (#PCDATA)>
]>
<name>
  <letter>J</letter>
  <digit>2</digit>
</name>
```

© Janusz R. Getta CSC1235/MCS9235/CSC1835 Databases, SCIT, Autumn 2015

15

03 Document Type Definition (DTD)

Attribute list

```
<!ATTLIST element-name attr-name attr-type attr-default-value >
```

Attribute types

CDATA	:any value is allowed (default)
(value ...)	:enumeration of allowed values
ID	:a value is a unique identifier
IDREF	:a value is an ID of another element (a reference to element)
IDREFS	:a value is a list of IDs of another element (the references to element)
NMTOKEN	:a value is a valid XML name
NMTOKENS	:a value is a list of valid XML names
ENTITY	:a value is an entity
ENTITIES	:a value is a list of entities
NOTATION	:a value is a name of a notation
xml:	:a value is a predefined XML value

© Janusz R. Getta CSC1235/MCS9235/CSC1835 Databases, SCIT, Autumn 2015

16

03 Document Type Definition (DTD)

Attribute list

```
<!ATTLIST element-name attr-name attr-type attr-default-value >
```

Attribute default values

value	:a given value is a default value
#REQUIRED	:a value is required
#IMPLIED	:a value is not required
#FIXED value	:a given value is fixed

© Janusz R. Getta CSC1235/MCS9235/CSC1835 Databases, SCIT, Autumn 2015

17

03 Document Type Definition (DTD)

Example of attribute list

Attribute type CDATA, default #REQUIRED

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE book [
  <!ELEMENT book (title)>
  <!ATTLIST book author CDATA #REQUIRED>
  <!ELEMENT title (#PCDATA)>
]>
<book author="C. J. Date">
  <title>Databases</title>
</book>
```

© Janusz R. Getta CSC1235/MCS9235/CSC1835 Databases, SCIT, Autumn 2015

18

03 Document Type Definition (DTD)

Example of attribute list

Attribute type ID, default #REQUIRED

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE book [
  <!ELEMENT book (title,chapter+)>
  <!ELEMENT title (#PCDATA)>
  <!ELEMENT chapter (#PCDATA)>
  <!--ATTLIST chapter id ID #REQUIRED-->
]>
<book>
  <title>Databases</title>
  <chapter id="I">Introduction</chapter>
  <chapter id="II">Conceptual modeling</chapter>
  <chapter id="III">SQL</chapter>
</book>
```

© Janusz R. Getta CSCI235/MCS9235/CSCI835 Databases, SCIT, Autumn 2015 19

03 Document Type Definition (DTD)

Example of attribute list

Attribute type (value|...), default #IMPLIED

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE degree [
  <!ELEMENT degree (subject+)>
  <!ELEMENT subject (#PCDATA)>
  <!--ATTLIST subject credits (6 | 12) #IMPLIED-->
]>
<degree>
  <subject credits="6">Databases</subject>
  <subject>Java for managers</subject>
  <subject credits="12">Software Project</subject>
</degree>
```

© Janusz R. Getta CSCI235/MCS9235/CSCI835 Databases, SCIT, Autumn 2015 20

03 Document Type Definition (DTD)

Example of attribute list

Attribute type CDATA, default #FIXED

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE book [
  <!ELEMENT book (title)>
  <!--ATTLIST book author CDATA #FIXED "J.R.Getta"-->
  <!--ELEMENT title (#PCDATA)-->
]>
<book author="C.J.Date">
  <title>Databases</title>
</book>
```

A document above does not validate !

Ln 7 Col 25 - Attribute "author" with value "C.J.Date" must have a value of "J.R.Getta".

© Janusz R. Getta CSCI235/MCS9235/CSCI835 Databases, SCIT, Autumn 2015 21

03 Document Type Definition (DTD)

Example of attribute list

Attribute type CDATA, default #FIXED

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE book [
  <!ELEMENT book (title)>
  <!--ATTLIST book author CDATA #FIXED "J.R.Getta"-->
  <!--ELEMENT title (#PCDATA)-->
]>
<book author="J.R.Getta">
  <title>Databases</title>
</book>
```

© Janusz R. Getta CSCI235/MCS9235/CSCI835 Databases, SCIT, Autumn 2015 22

03 Document Type Definition (DTD)

Example of attribute list

Attribute type ID, default #REQUIRED

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE book [
  <!ELEMENT book (chapter+)>
  <!ELEMENT chapter (#PCDATA)>
  <!--ATTLIST chapter number ID #REQUIRED-->
]>
<book>
  <chapter number="I">Introduction</chapter>
  <chapter number="I">Data Modeling</chapter>
  <chapter>Transaction processing</chapter>
</book>
```

A document above does not validate !

Ln 9 Col 22 - Attribute value "I" of type ID must be unique within the document.

Ln 10 Col 11 - Attribute "number" is required and must be specified for element type "chapter".

© Janusz R. Getta CSCI235/MCS9235/CSCI835 Databases, SCIT, Autumn 2015 23

03 Document Type Definition (DTD)

Example of attribute list

Attribute type ID, default #REQUIRED

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE book [
  <!ELEMENT book (chapter+)>
  <!ELEMENT chapter (#PCDATA)>
  <!--ATTLIST chapter number ID #REQUIRED-->
]>
<book>
  <chapter number="I">Introduction</chapter>
  <chapter number="II">Data Modeling</chapter>
  <chapter number="III">Transaction processing</chapter>
</book>
```

© Janusz R. Getta CSCI235/MCS9235/CSCI835 Databases, SCIT, Autumn 2015 24

03 Document Type Definition (DTD)
Example of attribute list
Attribute type IDREF, default #IMPLIED <pre><?xml version="1.0" encoding="UTF-8"?> <!DOCTYPE people [<!ELEMENT people (person*)> <!ELEMENT person (#PCDATA)> <!ATTLIST person ssno ID #REQUIRED> <!--ATTLIST person mother IDREF #IMPLIED--> <!--ATTLIST person father IDREF #IMPLIED--> <!--ATTLIST person child IDREF #IMPLIED-->]> <people> <person ssno="ssno-007" child="ssno-008">James Bond</person> <person ssno="ssno-001" child="ssno-008">Jane Moneypenny</person> <person ssno="ssno-008" father="ssno-007" mother="ssno-001"> James Bond Jr</person> </people></pre>
© Janusz R. Getta CSC1235/MCS9235/CSCI835 Databases, SCIT, Autumn 2015 25

03 Document Type Definition (DTD)
Example of attribute list
Attribute type IDREF, default #IMPLIED <pre><?xml version="1.0" encoding="UTF-8"?> <!DOCTYPE people [<!ELEMENT people (person*)> <!ELEMENT person (#PCDATA)> <!ATTLIST person ssno ID #REQUIRED> <!--ATTLIST person mother IDREF #IMPLIED--> <!--ATTLIST person father IDREF #IMPLIED--> <!--ATTLIST person child IDREF #IMPLIED-->]> <people> <person ssno="ssno-007" child="ssno-008">James Bond</person> <person ssno="ssno-001" child="ssno-008">Jane Moneypenny</person> <person ssno="ssno-008" father="ssno-006" mother="ssno-001"> James Bond Jr</person> </people></pre> <p>A document above does not validate ! Ln 15 Col 10 - An element with the identifier "ssno-006" must appear in the document.</p>
© Janusz R. Getta CSC1235/MCS9235/CSCI835 Databases, SCIT, Autumn 2015 26

03 Document Type Definition (DTD)
References
<p>Elmasri R., Navathe S., Fundamentals of Database Systems, 6th edition, chapter 12 XML: Extensible Markup Language, pp. 420-448 http://www.uow.edu.au/~jrg/235/HOMEWORK 8.2 How to create DTD and how to validate XML document against DTD ? (1) http://www.uow.edu.au/~jrg/235/HOMEWORK 8.3 How to create DTD and how to validate XML document against DTD ? (2) http://www.uow.edu.au/~jrg/235/SLIDES/XML/Appendix-1-DTD-Reference.pdf</p>
© Janusz R. Getta CSC1235/MCS9235/CSCI835 Databases, SCIT, Autumn 2015 27