1. **DTD trong XML**
2. Định nghĩa:

DTD – Document Type Definition, để xác định **cấu trúc** và sự hợp lý của các phần tử và thuộc tính trong tài liệu XML.

Chú ý: Tài liệu XML đúng cú pháp gọi là “**Well-Formed**”. Đúng cấu trúc gọi là “**validate**”.

1. Mục đích:

DTD trong XML sử dụng để **validate** tài liệu XML.

VD: Quy định tài liệu xml, phần tử root là danhsachsinhvien, và chỉ chứa 4 phần tử: id, name, age, address; Kiểu dữ liệu của 4 phần tử id, name, age, address là chuỗi và không chứa các thẻ con bên trong.

XML chỉ được phép thế này:

|  |
| --- |
| <?xml version="1.0" encoding="UTF-8"?>  <danhsachsinhvien>  <id>200</id>  <name>DungPT</name>  <age>25</age>  <address>Thanh Hóa</address>  </danhsachsinhvien> |
|  |

* Tài liệu XML kết hợp cùng DTD để giải quyết vấn đề trên.

1. Cú pháp DTD

* Tạo thẻ/ phần tử (Element) chỉ được phép chứa giá trị không chứa thẻ/ phần tử khác.

|  |
| --- |
| <!ELEMENT element-name (#PCDATA)> |

Trong đó: element-name là tên thẻ cần validate

VD: đối với thẻ id, name, age, address trên thể hiện:

<!ELEMENT id (#PCDATA)>

<!ELEMENT name (#PCDATA)>

<!ELEMENT age (#PCDATA)>

<!ELEMENT address (#PCDATA)>

* Tạo thẻ/ phần tử (Element) chứa phần tử khác

|  |
| --- |
| <!ELEMENT element-name (danh\_sách\_các\_phần\_tử\_con)> |

Trong đó: element-name là thẻ/ phần tử chứa các phần tử con.

danh\_sách\_các\_phần\_tử\_con Là danh sách các phần tử con của thẻ element-name, danh sách này là chuỗi ngăn cách nhau bởi dấu phẩy.

VD: thẻ danhsachsinhvien chứa 4 thẻ id, name, age, address

<!ELEMENT danhsachsinhvien (id,name,age,address)>

* Chỉ định một phần tử phải xuất hiện 1 lần

|  |
| --- |
| <!ELEMENT element-name (child-name)> |

Trong đó: element-name là thẻ/ phần tử chứa các thẻ con

child-name Là phần tử/ thẻ con bắt buộc phải xuất hiện 1 lần trong thẻ cha element-name

VD: Thẻ gender phải xuất hiện 1 lần trong thẻ danhsachsinhvien

<!ELEMENT danhsachsinhvien (gender)>

* Chỉ định một phần tử xuất hiện ít nhất 1 lần

|  |
| --- |
| <!ELEMENT element-name (child-name**+**)> |

Trong đó: element-name là thẻ/ phần tử chứa các thẻ con

child-name Là phần tử/ thẻ con bắt buộc phải xuất hiện ít nhất 1 lần trong thẻ cha element-name

VD: Thẻ note phải xuất hiện ít nhất 1 lần trong thẻ danhsachsinhvien

<!ELEMENT danhsachsinhvien (note**+**)>

* Chỉ định một phần tử không xuất hiện hoặc xuất hiện nhiều lần

|  |
| --- |
| <!ELEMENT element-name (child-name**\***)> |

Trong đó: element-name là thẻ/ phần tử chứa các thẻ con

child-name Là phần tử/ thẻ con có thể không xuất hiện hoặc xuất hiện nhiều lần trong thẻ cha element-name

VD: Thẻ note phải xuất hiện không xuất hiện hoặc xuất hiện nhiều lần trong thẻ danhsachsinhvien

<!ELEMENT danhsachsinhvien (note**\***)>

* Chỉ định 1 phần tử không xuất hiện hoặc xuất hiện 1 lần

|  |
| --- |
| <!ELEMENT element-name (child-name**?**)> |

Trong đó: element-name là thẻ/ phần tử chứa các thẻ con

child-name Là phần tử/ thẻ con có thể không xuất hiện hoặc xuất hiện 1 lần trong thẻ cha element-name

VD: Thẻ note không xuất hiện hoặc xuất hiện 1 lần trong thẻ danhsachsinhvien

<!ELEMENT danhsachsinhvien (note**?**)>

* Chỉ định một trong danh sách phần tử xuất hiện

|  |
| --- |
| <!ELEMENT element-name (child-name1|child-name2|child-name3)> |

Trong đó: element-name là thẻ/ phần tử chứa các thẻ con

child-name1, child-name2, child-name3 Là 1 trong 3phần tử/ thẻ con được xuất hiện trong thẻ cha element-name

VD: Thẻ note và thẻ identify là 1 trong 2 thẻ được xuất hiện trong thẻ danhsachsinhvien

<!ELEMENT danhsachsinhvien (id,name,age,address,(note|identify))>

* Khai báo thuộc tính cho phần tử

|  |
| --- |
| <!ATTLIST element-name **attribute-name** attribute-type attribute-value> |

Trong đó: element-name là thẻ/ phần tử có thuộc tính

**attribute-name là tên thuộc tính của thẻ**

attribute-type là kiểu thuộc tính, có thể là:

|  |  |
| --- | --- |
| **attribute-type** | **Mô tả** |
| CDATA | Giá trị là các ký tự |
| (*en1*|*en2*|..) | Giá trị phải là một từ một danh sách liệt kê |
| ID | Giá trị là id duy nhất |
| IDREF | Giá trị là id của thành phần khác |
| IDREFS | Giá trị là 1 list các ids khác |
| NMTOKEN | Giá trị là tên XML |
| NMTOKENS | Giá trị là list tên XML |
| ENTITY | Giá trị là 1 entity |
| ENTITIES | Giá trị là list các entities |
| NOTATION | Giá trị là tên notation |
| xml: | Giá trị là predefined xml value |

attribute-value là khai báo giá trị tương ứng của thuộc tính **attribute-name,** attribute-valuecó thể là:

|  |  |
| --- | --- |
| **attribute-value** | **Mô tả** |
| value | Giá trị mặc định của thuộc tính ( giá trị thực tế người dùng tạo ) |
| #REQUIRED | Giá trị của thuộc tính là bắt buộc |
| #IMPLIED | Giá trị của thuộc tính không bắt buộc |
| #FIXED value | Giá trị thuộc tính là cố định |

VD: Thẻ cha danhsachsinhvien chứa thẻ school có thuộc tính name = “FPT”:

<!ATTLIST school name CDATA "FPT">

Tương đương trong xml:

<school name="FPT"/>

1. Sử dụng trong XML document

* C1: Khai báo DTD ngay bên trong file XML:

|  |
| --- |
| <!DOCTYPE element DTD identifier  [  declaration1  declaration2  ........  ]> |

Trong đó:

* DTD bắt đầu với khai báo là <!DOCTYPE.
* Một element (phần tử/ thher) nói cho Parser để phân tích cú pháp tài liệu từ phần tử gốc đã cho.
* DTD identifier là một định danh cho Document Type Definition, mà có thể là path tới một file trên hệ thống hoặc URL tới một file trên mạng. Nếu DTD đang trỏ tới path bên ngoài, thì nó được gọi là External Subset.
* Dấu ngoặc móc vuông [ ] bao quanh một danh sách tùy ý các khai báo thực thể, được gọi là Internal Subset.
* File XML phải khai báo thêm thuộc tính standalone="yes"

VD:

|  |
| --- |
| <?xml version="1.0" encoding="UTF-8" standalone="yes" ?>  <!DOCTYPE address  [  <!ELEMENT address (name,company,phone)>  <!ELEMENT name (#PCDATA)>  <!ELEMENT company (#PCDATA)>  <!ELEMENT phone (#PCDATA)>  ]>  <address>  <name>DungPT</name>  <company>FPT</company>  <phone>09739647851</phone>  </address> |

* C2: Tạo file DTD riêng và khai báo trong file XML
* Thuộc tính standalone="no"
* Nghĩa là khai báo file DTD từ nguồn ngoài
* Cú pháp:

|  |
| --- |
| <!DOCTYPE root-element SYSTEM "file-name"> |

Trong đó:

root-element Là thành phần root của tài liệu XML

file-name Là đường dẫn tên file tới file .dtd

* Khi sử dụng file DTD ngoại vi bên ngoài có thể sử dụng định danh của hệ thống (System Identifier) hoặc định danh chung (Public Identifier)
* Định danh hệ thống (System Identifier): xác định vị trí của một file ngoại vi chứa các khai báo DTD
* Định danh chung (Public Identifier): xác định vị trí DTD resource

VD:

|  |
| --- |
| <?xml version="1.0" encoding="UTF-8"?>  <!ELEMENT sinhvien (id+,name,age,address,note\*,school,(identify|message))>  <!ELEMENT id (#PCDATA)>  <!ELEMENT name (#PCDATA)>  <!ELEMENT age (#PCDATA)>  <!ELEMENT address (#PCDATA)>  <!ELEMENT school (id, name)>  <!ELEMENT identify (#PCDATA)>  <!ELEMENT message (#PCDATA)>  <!ATTLIST school name CDATA #REQUIRED> |
| <?xml version="1.0" encoding="UTF-8"?>  <!DOCTYPE sinhvien SYSTEM "studentDTD.dtd">  <sinhvien>  <id>25</id>  <name>Dung</name>  <age>25</age>  <address>TH</address>  <school name="456">  <id>11</id>  <name>FPT</name>  </school>  <message>aa</message>  </sinhvien> |