

XML Technologies: Exercise Solutions 1

Introduction to XML, DTDs

Questions and instructions are displayed in a font with normal weight, answers are set in **bold**.

Bear in mind that even though the sample solution might not consider your approach to solve a particular problem, this does by no means invalidate your solution.

1 Introducing XML

1.1 Well-Formedness (2 Points)

Find all mistakes in the attempted XML document and correct them.

```
<?xml version="1.0" encoding="utf-8"?>
<!-- ? missing in declaration -->
<company>
  <person role="master distiller">
    <salary>10000</salary>
    <weight unit="kg">90</weight>
  </person>
  <!-- attributes must have unique names but tags do not -->
  <person>
    <roles>
      <role>CTO</role>
      <role>Business & Analyst</role>
      <!-- & must be escaped both in tags and attributes -->
      <role>CTO</role>
    </roles>
    <salary>9500</salary>
    <weight unit="kg">100</weight>
    <!-- watch out for typos! -->
    <!-- and that every attribute value is enclosed in quotes -->
  </person>
  <!-- and make sure every tag is closed -->
</company>
<!-- tags are closed with </tag>, not <tag/> -->
```

2 Validation and DTD

2.1 Validation and Valid Documents (1 Points)

- What is validation and what purpose does it serve?

Validation is the process of checking a well-formed XML document against a set of rules that are usually defined either in a DTD, XML Schema or RelaxNG. The rules all relate to the structure, ordering or content of XML.

- Under what conditions can an XML document be said to be valid?

An XML document is valid if it satisfies the rules defined in a DTD (or similar type of schema or definition language, for that matter).

- Can XML documents be well-formed, but not valid?

Yes, an absolute majority of all possible XML documents will be invalid with respect to a particular DTD at all times.

- Can XML documents be valid, but not well-formed?

No, an undefined something that is not well-formed XML cannot be validated.

2.2 Document Type Definitions (DTDs) (2 Points)

Write a DTD where the following rules are laid down:

- A `single_malt` element may only contain `whisky` or `whiskey` elements.
- Every `whisky` and `whiskey` element must have an `age` attribute.
- Every `whisky` and `whiskey` element may have an `origin` and `name` attribute.

```
<!ELEMENT single_malt (whisky|whiskey)+>
<!ELEMENT whisky EMPTY>
<!ELEMENT whiskey EMPTY>

<!ATTLIST whisky age CDATA #REQUIRED>
<!ATTLIST whisky origin CDATA #IMPLIED>
<!ATTLIST whisky name CDATA #IMPLIED>

<!ATTLIST whiskey age CDATA #REQUIRED>
<!ATTLIST whiskey origin CDATA #IMPLIED>
<!ATTLIST whiskey name CDATA #IMPLIED>
```