

Chapter 13

Question

Interchange
format



(Serialization)
n. 序列化, 串行化

• XML (Extensible Markup Language)

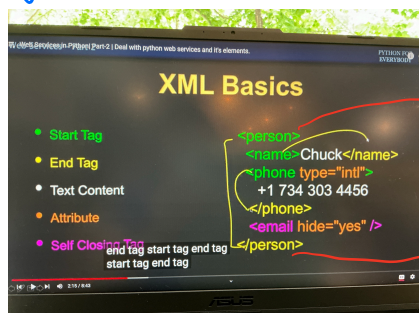
versus

- (1) Can both read by human and machine
- (2) XML is designed by human and aim to store and transfer data, emphasizing the tags of data

• HTML (HypeText Markup Language)

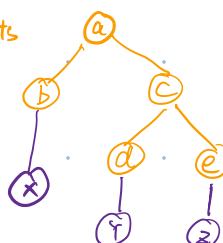
- (1) Can both read by human and machine
- (2) HTML is designed by human to the browser and aim to display contents on the web to viewer.

XML is like a Tree



Start with start tag
All include the tag (person) and All the data store here
End with End tag

—: Elements
—: Text



code:

```

<a>
  <b>X</b>
  <c>
    <d>Y</d>
    <e>Z</e>
  </c>
</a>
  
```

• XML Schema

Can check both way of data transformation, aiming to figure out exactly where the problem is.

XML documents

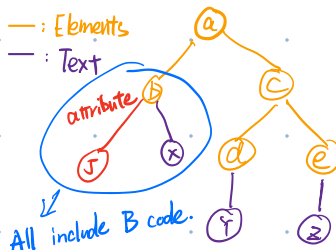
XML Validation



XML Schema Contract

Both of them should follow the rule and format of the Validator, and "documents" and "Contracts" need to match each other.

—: Elements
—: Text



code

```

<a>
  <b w="r">X</b>
  <c>
    <d>Y</d>
    <e>Z</e>
  </c>
</a>
  
```

• Usages in Python

- (1) import xml.etree.ElementTree as ET
- (2) Triple quoted means multi-line string

example

```

data = '''<person>
  <name>Chuck</name>
  <phone type="intl">
    +1 734 303 4456
  </phone>
  <email hide="yes"/>
</person>'''
  
```