

Introduction to XML

eXtensible
Markup
Language

10011110101001



bits and bytes

ABCDEF



ASCII

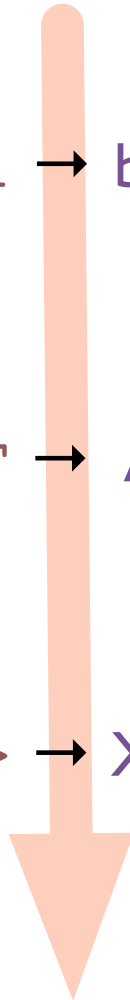
(alphabet)

<tag/>

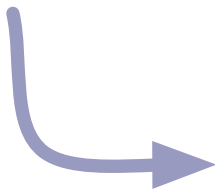


XML

(punctuation)



SGML (Standard Generalized Markup Language)

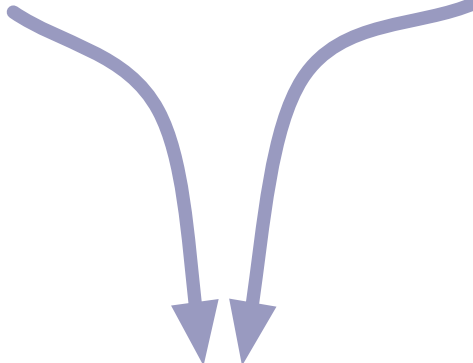


HTML

(HyperText Markup Language)

CML

(Chemical Markup Language)



XML

XML rules (ok?)

5 things you need to know:

Elements
Attributes
Text
Escaping
Comments

```
<?xml version="1.0" encoding="utf-8"  
standalone="no"?>
```

XML declaration

<?PITarget PIdata?>

Processing Instruction

```
<!DOCTYPE name [ <!MORE RANDOM STUFF> ]>
```

DocType Declaration

Elements

```
<parent>  
  <child>  
    <grandchild/>  
  </child>  
</parent>
```

Elements

```
<parent>  
  <child>  
    <grandchild/>  
  </child>  
</parent>
```

Opening Tags

Elements

```
<parent>  
  <child>  
    <grandchild/>  
  </child>  
</parent>
```

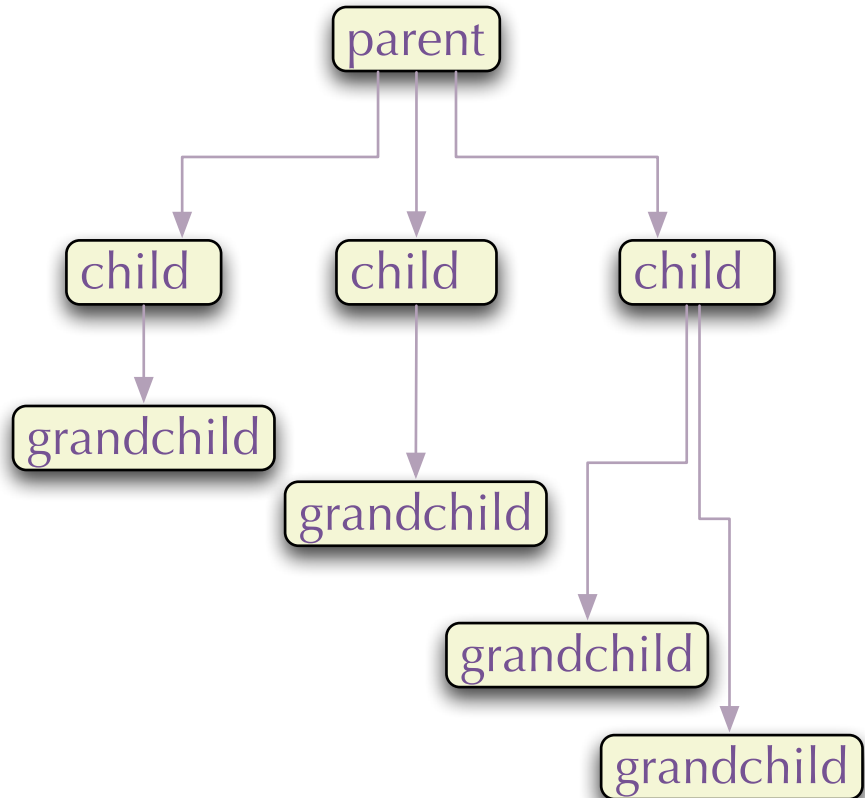
Closing Tags

Elements

```
<parent>  
  <child>  
    <grandchild/>  
  </child>  
</parent>
```

Empty Tag

```
<parent>  
  <child>  
    <grandchild/>  
  </child>  
  <child>  
    <grandchild/>  
  </child>  
  <child>  
    <grandchild/>  
    <grandchild/>  
  </child>  
</parent>
```



**XML trees:
Elements**

XML names

(Letters, _) Numbers, _, .

Case sensitive

thisIsAName

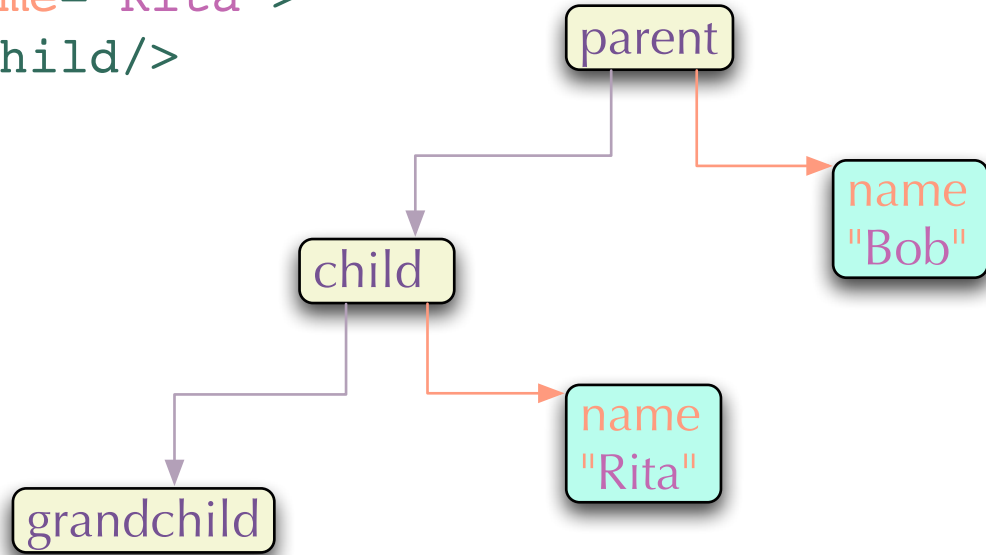
2thisIsNot

this \neq This

Attributes

```
<parent name="Bob">  
  <child name="Verity" nickname="Rita">  
    <grandchild name=' '  
      nickname='baby' />  
  </child>  
</parent>
```

```
<parent name="Bob">  
  <child name="Rita">  
    <grandchild/>  
  <child/>  
</parent>
```

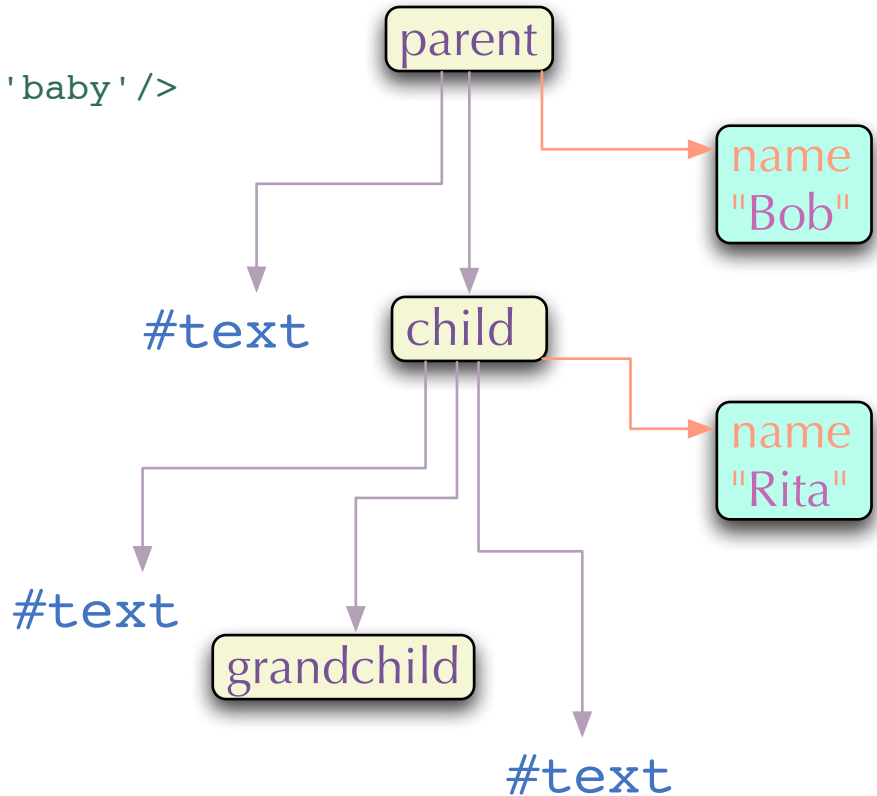


**XML trees:
attribute names and values**

Text

```
<parent name="Bob">
  Bob is a doctor
  <child name="Verity" nickname="Rita">
    Rita is a lawyer.
    <grandchild name=' '
      nickname='baby' />
    Rita is Bob's child.
  </child>
</parent>
```

```
<parent name="Bob">
  Bob is a doctor
  <child name="Verity" nickname="Rita">
    Rita is a lawyer.
    <grandchild name=' '
      nickname='baby' />
    Rita is Bob's child.
  </child>
</parent>
```



XML trees:
text

Escaping

```
<element quotemark="&quot;">  
  A left angle bracket is &lt;  
  A right angle bracket is &gt;  
  An apostrophe is &apos;  
  A quotation mark is &quot;  
  And an ampersand is &amp;  
</element>
```

Comments

```
<!-- This is a comment -->
```

But what if I
get it wrong?