

Структура документов XML

Если использовать классические термины программирования, то документ XML наиболее близок к такой структуре данных, как дерево. Дерево представляет собой упорядоченный набор связанных элементов, называемых узлами. Принципы организации узлов в дерево довольно просты. Каждый узел связан с одним узлом, называемым родительским. Каждый узел может иметь любое число дочерних узлов, для которых он сам будет родительским узлом. Лежащий в основании дерева узел не имеет родительского узла и называется корневым узлом. Уровнем узла дерева называется число, определяющее, сколько элементов нужно пройти, чтобы добраться до корневого узла. У узла могут быть индивидуальные свойства, называемые атрибутами. Наиболее очевидный пример дерева — это SWF-фильм. Узлами в нем можно считать клипы. Корневой узел — это основная временная диаграмма. У нее нет родительского узла, так как она не относится ни к одному клипу. Клипы, располагающиеся на `_root`, являются узлами первого уровня. Вложенные в них клипы — это узлы второго уровня (и так далее). У каждого узла есть только один родительский узел (клип может относиться только к одной временной диаграмме), но может быть много дочерних узлов (в клип можно вложить любое число клипов). Клип связан с родительским клипом свойством `_parent`, а с дочерними клипами — свойствами, имеющими такие же имена, как и у них. У клипа есть атрибуты — его собственные свойства и методы. Узел в XML принято называть элементом. Однако мы все же будем использовать термин “узел”, так как он лучше характеризует древовидную организацию XML-документа. В XML узлы задаются при помощи тегов. Простейший XML-документ может быть образован одним лишь корневым узлом:

```
<name>Jonh Smith</name>
```

Все, что располагается между тегами, задающими узел, называется содержимым узла. Возможно два типа содержимого: текст и другие узлы (содержимое может быть смешанным). Вкладывая в тег теги, вы создаете дочерние узлы. Соответственно, тег узла, содержащий тег данного узла, является для него родительским. Вложенный в тег текст также можно представлять, как дочерний узел (именно так действует анализатор XML Flash-плеера). Количество дочерних узлов, а также уровней в дереве не лимитировано. Совершенно очевидно, почему XML-документ имеет структуру дерева. С одной стороны, тег может быть вложен только в один тег. С другой стороны, в тег можно вложить любое число тегов (проведите аналогию с клипами SWF-фильма). Для примера приведем многоуровневый XML-документ, в котором хранятся анкетные данные гипотетического американца:

<person>

<name>Jonh Smith</name>

<age>44</age>

<profession>manager</profession>

<family>

<wife>

<name>Mary</name>

<age>39</age>

<profession>teacher</proffesion>

</wife>

<child>

<name>Bob</name>

<age>18</age>

<profession>student</profession>

</child>

</family>

</person>

Объективно говоря, приведенная диаграмма верна не до конца. Для узла дерева важно не только количество дочерних узлов, но и порядок их следования. Чем раньше располагается соответствующий узлу тег, тем левее он должен быть отображен на диаграмме. Несложно заметить, что диаграмма 1 не выполняет это требование. Очень важно научиться представлять XML-документ как дерево. Дело в том, что при разборе документа анализатором Flash-плеера, формируется реальное дерево из объектов классов XML и XMLNode. И чтение данных из XML, а также преобразование документа основывается на обходе этого дерева.