### Лабораторная работа № 18

**Тема:** Разработка программ создания и обработки XML-документов

**Цель работы:** Формирование умений и навыков создания и обработки XML-документов

Для чтения **XML** чаще всего используются высокоуровневые классы **XmlDocument** и **XPathDocument**. Что не удивительно, ведь ими довольно легко пользоваться. Данные представлены в логичном, понятном виде, можно в любой момент получить любой набор данных в нужно формате. Достигается это за счет того, что xml-данные загружаются в память. Это некритично для сравнительно небольших xml-файлов. Однако когда размер xml-данных доходит до нескольких сотен мегабайт, а то и гигабайт, классы **XmlDocument** или**XPathDocument** использовать просто неприлично, а в некоторых случаях и вовсе невозможно, ведь приложение будет требовать объем оперативной памяти равный размеру данных. Именно для таких случаев и предназначен класс **XmlReader**.

В отличие ОТ своих собратьев ПО пространству имен, класс XmlReader не данные В памяти. Доступ хранит данным осуществляется последовательно. Но из-за последовательности становится невозможным узнать, что содержит тот или иной xml-элемент, до того, как будут прочитаны. Этот недостаток данные делает класс XmlReader неудобным в использовании и даже может отпугнуть неопытных программистов. Но все не так страшно, как может показаться на первый взгляд.

### Пример чтения ХМL

Чтобы было понятно, о чем идет речь, приведу простой пример. Предположим, что у нас есть xml-файл с информацией о пользователях:

Данные отдельного пользователя находятся в элементе member и содержат: идентификатор пользователя (атрибут kuid), псевдоним (элемент nickname), имя (firstName) и фамилию (lastName).

```
<?xml version="1.0" encoding="utf-8" ?>
<kbyte>
 <members>
    <member kuid="1">
      <nickname>Алексей</nickname>
      <firstName>Алексей</firstName>
     <lastName>Heмиро</lastName>
   </member>
    <member kuid="288">
      <nickname>Игорь Голов</nickname>
     <firstName>Игорь</firstName>
      <lastName>Голов</lastName>
    </member>
    <member kuid="1858">
      <nickname>[i]Pro</nickname>
      <firstName>Aprem</firstName>
      <lastName>Донцов</lastName>
   </member>
    <member kuid="2575">
      <nickname>Shark1</nickname>
      <firstName>Виталий</firstName>
      <lastName />
   </member>
  </members>
</kbyte>
```

При помощи класса **XmlDocument** данные каждого пользователя можно получить следующим образом:

**Примечание.** Классы **XmlDocument** и **XmlNode** принадлежат пространству имен **System.Xml**. Для их использования необходимо импортировать это пространство в проект.

**Примечание.** Предполагается наличие файла с именем **XMLFile1.xml** в корневом каталоге приложения, который содержит приведенные выше xml-данные.

При использовании класса **XPathDocument** принципиальных отличий в коде не будет.

**Примечание.** Классы **XPathDocument** и **XPathNavigator** принадлежат пространству имен **System.Xml.XPath**, которое необходимо импортировать в проект.

Класс **XPathDocument** отличается от **XmlDocument** в основном тем, что позволяет использовать механизмы языка **XPath** (**XML Path Language**). Язык **XPath** содержит множество встроенных функций для обработки информации, но в рамках данной статьи эта тема рассматриваться не будет.

Как видино из вышеприведенных примеров, при использовании классов **XmlDocument** и **XPathDocument** доступ к данным осуществляется через коллекцию узлов (**node**), каждый из которых имеет функции и методы управления содержимым документа. В случае с классом**XmlReader** все будет менее очевидно, т.к. доступ к данным осуществляется на более низком уровне.

XmlReader читает данные последовательно ИЗ потока. Свойство **NodeType** содержит информацию о типе прочитанных данных. При последовательном чтении данных, невозможно сразу узел /kbyte/members/member. До него должен дойти указатель. Но и после того, как узел будет найден, невозможно узнать, какие данные он содержит, пока указатель не дойдет до этих данных. Код получается сложным для восприятия. И это притом, что в представленном примере структура xmlдокумента простая. Обработка документов со сложной структурой потребует написания более мудреного кода. Хотя, если хорошенько подумать, даже в самых тяжелых случаях можно упростить задачу. Об этом пойдет речь далее.

### Работа с XmlReader

Класс **XmlReader** не имеет конструктора. Создать новый экземпляр класса **XmlReader** можно при помощи метода **Create**, который принимает либо физический путь или **url** к xml-документу, либо поток (**Stream**).

```
XmlReader xml = XmlReader.Create("XmlFile1.xml");
```

Для перехода от одного узла к другому предназначена функция Read. которая возвращает либо **true**, если удалось осуществить переход К узлу, либо, ПО достижению конца документа, false. Как правило, чтение всего документа производится циклом.

```
while (xml.Read())
{
// ..
}
```

В зависимости от типа узла (**NodeType**), экземпляр класса **XmlReader** может содержать дополнительно имя узла (**Name**) и его текстовое значение (**Value**).

```
Console.WriteLine($"Тип узла: {xml.NodeType}");;
Console.WriteLine($"Имя узла: {xml.Name}");
Console.WriteLine($"Значение узла: {xml.Value}");
```

Для проверки наличия данных, предназначены вспомогательные свойства: **HasValue** — указывает, имеет узел текстовое значение (**Value**) или нет; **HasAttributes** — указывает, имеет узел атрибуты или нет; **IsEmptyElement** — содержит **true**, если элемент не имеет никаких данных.

Если текущий узел содержит атрибуты (**HasAttributes** = **true**), то получить их можно при помощи функции **MoveToNextAttribute**, которая, по аналогии с **Read**, будет возвращать **true**, до достижения конца определения элемента. Имя атрибута и значение также будут находиться в свойствах **Name** и **Value**.

```
if (xml.HasAttributes)
{
    while (xml.MoveToNextAttribute())
    {
        Console.WriteLine($"Атрибут: {xml.Name}");
        Console.WriteLine($"Значение: {xml.Value}");
    }
}
```

После считывания атрибутов может потребоваться вернуться назад к элементу, для этого существует метод **MoveToElement**.

```
if (xml.HasAttributes)
{
    while (xml.MoveToNextAttribute())
    {
        Console.WriteLine($"Атрибут: {xml.Name}");
        xml.MoveToElement();
        Console.WriteLine($"Значение: {xml.Value}");
    }
}
```

Для упрощения обработки документа есть несколько полезных функций. Методы **ReadOuterXml** и **ReadInnerXml** позволяют получить фрагменты **xml** в виде текста, как есть. Так, например, найдя в документе элемент **member**, при помощи функции **ReadOuterXml** можно получить его целиком и загрузить в **XmlDocument**, т.е. обработать данные на более высоком уровне, а не собирать их по крупинкам, как в продемонстрированном ранее примере.

А при желании, можно вообще сделать класс **Member** и создать его экземпляр на основе полученной из **xml** информации при помощи механизмов десериализации. **NET Framework**.

```
[XmlRoot("member")]
public class Member
    [XmlAttribute("kuid")]
   public int KUID { get; set; }
    [XmlElement("nickname")]
    public string Nickname { get; set; }
    [XmlElement("firstName")]
    public string FirstName { get; set; }
    [XmlElement("lastName")]
   Ссылок: 0
    public string LastName { get; set; }
    public Member() { }
   Ссылок: 0
    public Member(string xml)
        LoadXml(xml);
    public void LoadXml(string source)
        XmlSerializer mySerializer = new XmlSerializer(this.GetType());
        using (MemoryStream ms = new MemoryStream(Encoding.UTF8.GetBytes(source)))
            object obj = mySerializer.Deserialize(ms);
            foreach (PropertyInfo p in obj.GetType().GetProperties())
                PropertyInfo p2 = this.GetType().GetProperty(p.Name);
                if (p2 != null && p2.CanWrite)
                    p2.SetValue(this, p.GetValue(obj, null), null);
```

**Примечание.** Для использования классов xml-сериализации необходимо импортировать в проект пространство имен**System.Xml.Serialization**. Помимо этого, для работы функции **LoadXml** может потребоваться импортировать пространства имен**System.IO** и **System.Reflection**.

При этом уровень обработки данных повышается не для всего содержимого **xml**, объем которого может достигать многих гигабайт, а лишь для его небольшого фрагмента. Потребление памяти, да и в целом нагрузка на компьютер, незначительные. А вот что делать, если элемент **member** будет содержать большие объемы вложенной информации. Например, все сообщения пользователя с форумов **Kbyte.Ru**?

```
<?xml version="1.0" encoding="utf-8" ?>
∃<kbyte>
  <members>
     <member kuid="1">
       <nickname>Алексей</nickname>
       <firstName>Алексей</firstName>
       <lastName>Heмиро</lastName>
       <messages>
         <message id="1">
           <subject>Привет мир!</subject>
           <text>Это тестовое сообщение.</text>
         </message>
         <message id="2">
           <subject>Как работать с XmlReader?</subject>
           <text>Это текст о том, как работать с XmlReader.</text>
         </message>
       </messages>
     </member>
     <member kuid="288">
       <nickname>Игорь Голов</nickname>
       <firstName>Игорь</firstName>
       <lastName>Голов</lastName>
     </member>
    <member kuid="1858">
       <nickname>[i]Pro</nickname>
       <firstName>Aprem</firstName>
       <lastName>Донцов</lastName>
       <messages>
         <message id="3">
           <subject>Eще одно сообщение</subject>
           <text>Это текст еще одного тестового сообщения.</text>
         </message>
       </messages>
     </member>
     <member kuid="2575">
       <nickname>Shark1</nickname>
       <firstName>Виталий</firstName>
       <lastName />
       <messages />
     </member>
   </members>
```

В этом документе, каждый элемент member содержит один вложенный элемент **messages**, который в свою очередь может состоять из неограниченного числа элементов **message**. Если получать все содержимое узла **member**, то потребление оперативной памяти может быть высоким.

Поэтому использование функции **ReadOuterXml** для этих целей не годится. В подобных случаях нужно разделить логику обработки элементов **member** и **messages**. Таким образом, чтобы элемент **member** содержал только основные данные о пользователе, как и в первых примерах, исключив из него элемент **messages**. Но и список сообщений никуда не пропадает, он просто будет обработан отдельно, причем ничто не запрещает использовать для этого метод **ReadOuterXml**.

Чтобы исключить объемные фрагменты **xml**, необходимо записывать в память только то, что нужно. Для этого потребуется создать экземпляр класса **XmlWriter**. Запись данных проще всего производить в **StringBuilder**, хотя можно и в поток (**Stream**). Также потребуется две дополнительных булевых (логических) переменных, одна из которых будет разрешать запись данных в экземпляр **XmlWriter**, а вторая — информировать об исключении **xml** элемента.

```
StringBuilder sb = null;
XmlWriter w = null;
XmlWriterSettings ws = new XmlWriterSettings() { Encoding = Encoding.UTF8 };
bool isMember = false, isSkiped = false;
```

Пропустить ненужный фрагмент **xml** можно при помощи метода **Skip**.

```
using (XmlReader xml = XmlReader.Create("XmlFile1.xml"))
    while (xml.Read())
        switch (xml.NodeType)
            case XmlNodeType.Element:
                 if (xml.Name == "member")
                     isMember = true;
                     if (sb != null)
                         w.Flush();
                         w.Close();
                         XmlDocument xmlDoc = new XmlDocument();
                         xmlDoc.LoadXml(sb.ToString());
                         XmlNode n = xmlDoc.SelectSingleNode("member");
                         Console.WriteLine("KUID: {0}", n.Attributes["kuid"].Value);
                         Console.WriteLine("Псевдоним: {0}", n.SelectSingleNode("nickname").InnerText);
                         Console.WriteLine("Имя: {0}", n.SelectSingleNode("firstName").InnerText);
Console.WriteLine("Фамилия: {0}", n.SelectSingleNode("lastName").InnerText);
                         Console.WriteLine("-----
                     // создаем новый XmlWirter, для записи данных пользователя
                     sb = new StringBuilder();
                     w = XmlWriter.Create(sb, ws);
                     w.WriteProcessingInstruction("xml", "version=\"1.0\" encoding=\"utf-8\"");
                 else if (xml.Name == "messages")
                     xml.Skip();
                     isSkiped = true;
```

```
// это данные пользователя и не пропущенный фрагмент
    if (isMember && !isSkiped)
       w.WriteStartElement(xml.Name);
        // если есть атрибуты, записываем их
        if (xml.HasAttributes)
           while (xml.MoveToNextAttribute())
                w.WriteAttributeString(xml.Name, xml.Value);
    // убираем отметку о пропущенном фрагменте
    isSkiped = false;
    break;
case XmlNodeType.Text:
    if (isMember)
       w.WriteString(xml.Value);
    break;
case XmlNodeType.EndElement:
    if (isMember)
    {
       w.WriteFullEndElement();
    if (xml.Name == "member")
        isMember = false;
    break;
```

По функционалу приведенный листинг ничем не отличается от предыдущих. На выходе также будет получен фрагмент документа, содержащий только элемент **member**, исключая вложенный элемент **messages**. Что касается обработки **messages**, то для этого проще всего сделать отдельный **XmlReader** при помощи функции **ReadSubtree**.

```
else if (xml.Name == "messages")
   XmlReader xml2 = xml.ReadSubtree();
   while (xml2.Read())
       switch (xml2.NodeType)
           case XmlNodeType.Element:
              if (xml2.Name == "message")
                  XmlDocument xmlDoc2 = new XmlDocument();
                  xmlDoc2.LoadXml(xml2.ReadOuterXml());
                  XmlNode n = xmlDoc2.SelectSingleNode("message");
                  Console.WriteLine("Сообщение # {0}", n.Attributes["id"].Value);
                  Console.WriteLine("Tema: {0}", n.SelectSingleNode("subject").InnerText);
                  Console.WriteLine("Texcr: {0}", n.SelectSingleNode("text").InnerText);
                  Console.WriteLine("-----");
              break;
   // пропускаем фрагмент с сообщениями
   xml.Skip();
   isSkiped = true;
```

Таким образом, у нас будет отдельно обработан каждый элемент **member** и каждый связанный с ним **message**. Это практически никак не отразится на потреблении ресурсов, каким бы большим не был размер xmlфайла.

По завершению работы с документом, всегда необходимо закрывать поток при помощи метода **Close**.

**Примечание.** При использовании инструкции **using** { }, закрытие потока происходит автоматически. Отдельный вызов метода **Close** может привести к возникновению исключения.

Использование сериализации для сохранения в xml.

```
using System.Xml.Serialization;
using System.IO;

namespace XMLserial
{
    [Serializable]
    ccылок 6
    public class Example
    {
        ccылка 1
        public string Name { set; get; }
        ccылка 1
        public uint Number { set; get; }
        ccылок 0
        public Example() { }
        ccылка 1
        public Example(string name, uint number)
        {
              this.Name = name;
              this.Number = number;
        }
    }
}
```

```
Static void Main(string[] args)

{

Example[] mas = new Example[10];

Random rnd = new Random();

uint number = 1;

for(int i = 0; i<mas.Length;i++)

{

    mas[i] = new Example("Example" + i,number);

    number += (uint)rnd.Next(0, 10);

}

XmlSerializer formatter = new XmlSerializer(typeof(Example[]));

//coxpahutb

using(StreamWriter sw = new StreamWriter("test.xml"))

{

    formatter.Serialize(sw, mas);

}

//sarpysutb

using (StreamReader sr = new StreamReader("test.xml"))

{

    Example[] newmas = (Example[])formatter.Deserialize(sr);

}
```

```
XmlDocument xdock = new XmlDocument();
xdock.Load("test.xml");
foreach (XmlNode el in xdock.SelectNodes("//Example"))
{
    string Number = el.SelectSingleNode("Number").InnerText;
    int nub = Convert.ToInt32(Number);
    if (nub>5 && nub<20)
    {
        Console.Write("Упражнение " + el.SelectSingleNode("Name").InnerText);
        Console.WriteLine(" № {0} ", el.SelectSingleNode("Number").InnerText);
    }
}</pre>
```

# Практическая часть

# Задание1

Разработать структуру из варианта. Сохранить и загрузить файл. Реализовать операцию из варианта

Вариант	Задание
1	Структура «Автосервис»: регистрационный номер автомобиля, марка, пробег,
	мастер, выполнивший ремонт, сумма ремонта.
	Операция: удаление каждого 2го элемента, поиск по марке автомобиля
2	Структура «Сотрудник»: фамилия, имя, отчество; должность; год рождения;
	заработная плата.
	Операция: запись нового элемента после заданного, поиск по зарплате больше
	заданной.
3	Структура «Государство»: название; столица; численность населения; занимаемая
	площадь.
	<b>Операция</b> : удаление последнего элемента, поиск по численности в границе([a,b]),
	где а,b – вводятся с клавиатуры.
4	Структура «Человек»: фамилия, имя, отчество; домашний адрес; номер телефона;
	возраст.
	Операция: удаление предпоследнего элемента, поиск по номеру телефона зная
	первые 3 цифры(из 7).
5	Структура «Читатель»: Фамилия И.О., номер читательского билета, название
	книги, срок возврата. Операция: запись нового элемента в середину файла, поиск по номеру
	читательского билета.
6	Структура «Школьник»: фамилия, имя, отчество; класс; номер телефона; оценки
	по предметам (математика, физика, русский язык, литература).
	Операция: запись нового элемента в середину файла, поиск по номеру телефона.
7	Структура «Студент»: фамилия, имя, отчество; домашний адрес; группа; рейтинг.
	Операция: удаление 1го элемента, поиск по фамилии.
8	Структура «Покупатель»: фамилия, имя, отчество; домашний адрес; номер
	телефона; номер кредитной карточки.
	<b>Операция</b> : запись нового элемента в конец, поиск по номеру кредитной карты известны 4 из 16дцати цифр.
9	известны 4 из годцати цифр.  Структура «Пациент»: фамилия, имя, отчество; домашний адрес; номер
	медицинской карты; номер страхового полиса.
	Операция: запись нового элемента в конец, поиск по номеру страхового полиса
	известны 4.
10	Структура «Информация»: носитель; объем; название; автор.
	Операция: удаление каждого 3го элемента, поиск по автору.
11	Структура «Клиент банка»: Фамилия И.О., номер счета, сумма на счете, дата
	последнего изменения.
	Операция: запись после заданного, поиск сумме больше заданной.
12	Структура «Склад»: наименование товара, цена, количество, процент торговой
	надбавки.
	Операция: удаление первого и последнего элемента, поиск по количеству больше
	заданного для произвольного поля.
13	Структура «Авиарейсы»: номер рейса, пункт назначения, время вылета, дата
	вылета, стоимость билета. Операция: удаление каждого 2го элемента, поиск по номеру рейса.
	Операция. удаление каждого 210 элемента, поиск по номеру реиса.

## Задание 2.

Разработка программ создания и обработки XML-документов

#### Общее задание:

Вам необходимо разработать магазин, ведущий учет своих товаров через XML файл. У каждого товара есть наименование и не менее 2х характеристик, а также количество на складе. Необходимо написать приложение в котором будет реализовано:

- 1. Поставка (существующие модели увеличиваются в количестве),
- 2. Покупка (только при наличии),
- 3. Поиск по обеим характеристикам (на точное совпадение)
- 4. Действие в варианте.

Использовать только XML.Linq.

#### Варианты

- 1. Магазин телефонов. Действие: удаления со склада телефонов, если их количество 0.
- 2. Магазин телевизоров. Действие: создать клон заданной модели
- 3. Магазин компьютеров. Действие: поиск по названию. Выводить все модели которые в названии встречают строку поиска, указанную пользователем.
- 4. Магазин игрушек. Действие: удаление после указанной модели.
- 5. Магазин цветов. Действие: вставка перед указанной моделью.
- 6. Магазин продуктов. Действие: удаление перед указанной моделью.
- 7. Магазин канцтоваров. Действие: вставка после указанной моделью.
- 8. Магазин парфюмерии. Действие: антипоиск по названию. Выводить все модели которые в названии **не** встречают строку поиска, указанную пользователем.

## Контрольные вопросы:

- 1. Особенности разметки XML документа.
- 2. Класс XmlNode.
- 3. Класс XmlNodeList.
- 4. Что такое XPath. И основные запросы
- 5. Класс XmlReader

#### Литература

- 1. Полный справочник по С#. Г. Шилдт. Издательский дом «Вильямс», 2004.
- 2. С# в подлиннике. Наиболее полное руководство. Х.Дейтел.

- 3.
- 4.
- С# в задачах и примерах. Культин. Н.Б. С# учебный курс. Г.Шилдт. СПб.: Питер, 2002. С# программирование на языке высокого уровня Павловская Т.А. СПб.: 5.

БХВ-Петербург.