

## АЛГОРИТМ СТРУКТУРОВАНОЇ ВІЗУАЛІЗАЦІЇ XML-ФАЙЛІВ

Трояновська Тетяна, Бойко Олександр

Вінницький національний технічний університет

### Анотація

Дана робота присвячена розробці засобу для перегляду файлів у форматі XML, який за останні роки став одним із базових носіїв даних практично у всіх галузях ІТ-індустрії, включаючи ігрову сферу та бізнес-логіку. Метою роботи є розробка засобу, який дозволяв би перегляд змісту XML-файлу в інтерактивній формі – це пришвидшить пошук необхідного фрагменту. Високий рівень виконання поставленої задачі вирішено засобами мови програмування Java.

### Abstract

This work is devoted to the development of programs to view files, compiled in XML format, which in recent years has become one of the basic media data in almost all areas of the it industry, including the gaming sphere and the business logic. However, in the case of large XML files, it is advisable to have a software tool that would allow viewing of its content in an interactive form - this will speed up the search for a desired fragment. The high level task solving by means of the Java programming language is achieved.

### Вступ

XML (Extensible Markup Language) - це мова розмітки, що описує цілий клас об'єктів даних, які називаються XML-документами і використовується для опису граматики інших мов і контролю за правильністю складання документів. Сам по собі XML не містить тегів, призначених для розмітки, він визначає порядок їх створення і дозволяє здійснювати контроль за коректністю даних, робити перевірки ієрархічних співвідношень усередині документа, встановлювати єдиний стандарт на структуру документів. Це означає, що XML можна використовувати при побудові складних інформаційних систем, де важливим є питання обміну інформацією між різними додатками, що працюють в одній системі [1, 2].

### Основний виклад матеріалу

Візуалізація XML-файлу може бути описана таким базовим алгоритмом: 1) Аналіз XML-документа, перевірка на правильність та синтаксичну коректність; 2) Визначення кореневого елемента; 3) Рекурсивне завантаження всіх елементів, що містяться в кореновому; 4) Відображення елементів у вигляді деревовидної структури.

Всі необхідні для реалізації цього алгоритму класи та компоненти присутні в стандартній поставці мови програмування Java і не вимагають сторонніх бібліотек.

Розглянемо детально кожен із описаних нами кроків візуалізації XML-документа.

1) *Аналіз розробленого XML – документа та перевірка його на правильність.* Для цього використовується компонент JAXP – Java API for XML Processing, що містить два парсери XML-файлів: SAX – Simple API for XML та DOM – Document Object Model (рис.2).

Як показано на рис.1, в рамках аналізатору SAX є 4 типи подій, що виникають при аналізі XML – документа: а) контент (під цим терміном мається на увазі як елементи, так і їх контент); б) помилка розбору; в) схема документу; г) спеціальні символні комбінації.

Такий підхід дозволяє виокремити і обробити фрагмент XML-документу, призначивши відповідний обробник і відстеживши відкриття потрібного тегу [3]. Натомість ядро DOM (рис.2) являє собою деревовидну структуру XML-документу, і його аналізатор завантажує XML-документ в неї, таким чином полегшуючи як перевірку на

коректність, так і роботу із документом [4]. Зауважимо, що цей аналізатор виконує 3 перших пункти алгоритму візуалізації.

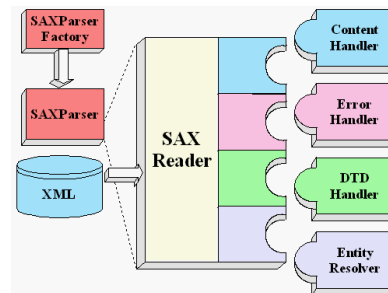


Рисунок 1 – SAX-аналізатор

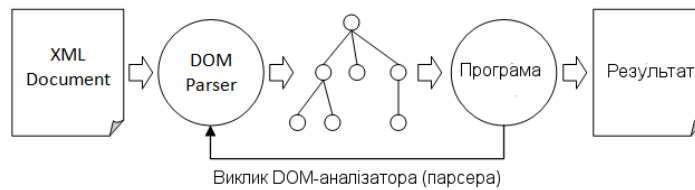


Рисунок 2 – DOM-аналізатор

Код для застосування DOM-аналізатора наведений у лістингу 1.

```

FileInputStream fis = new FileInputStream(fileName);
DocumentBuilderFactory factory = DocumentBuilderFactory.newInstance();
DocumentBuilder builder = factory.newDocumentBuilder();
Document document = builder.parse(fis);
    
```

Лістинг 1 – Аналіз та завантаження змісту XML-документа

2) *Визначення кореневого елемента.* Для візуалізації змісту необхідне деревовидне представлення, що створюється за допомогою рекурсивного перегляду і перенесення відповідних даних в елемент керування: знайти елемент керування, який здатен відобразити дерево і створити таблицю переліку атрибутів та їх значень. Такі елементи містяться в Java, її графічній бібліотеці Swing, що входить в стандартну поставку.

3) *Рекурсивне завантаження всіх елементів документа.* Swing – один із найбільш комплексних компонентів мови програмування Java. Він містить повний набір елементів керування для конструювання графічної програми будь-якого призначення.

4) *Відображення елементів у вигляді деревовидної структури.* Для візуалізації моделі у вигляді деревовидної структури Document нам знадобляться компоненти JTree та JTable [5].

Елемент дерева є компонентом-контейнером DefaultMutableTreeNode. Обійдемо дерево, і встановимо відповідність між елементом дерева та компонентом (лістинг 2).

```

private void traverseXML(Element root,
    DefaultMutableTreeNode treeRoot)
{
    if(root.getChildNodes())
    {
        NodeList subs = root.getChildNodes();
        for(int i = 0; i < subs.getLength(); i++)    if (subs.item(i) instanceof Element)
        {
            DefaultMutableTreeNode newNode = new
            DefaultMutableTreeNode(new
            XMLTreeNode((Element)subs.item(i)));
            traverseXML((Element)subs.item(i), newNode);
            treeRoot.add(newNode); }
    }
    
```

Лістинг 2.4 – Рекурсивне співставлення DOM-дерева дереву елемента керування

Таким чином, описаний алгоритм структурованої візуалізації реалізований (рис. 3).

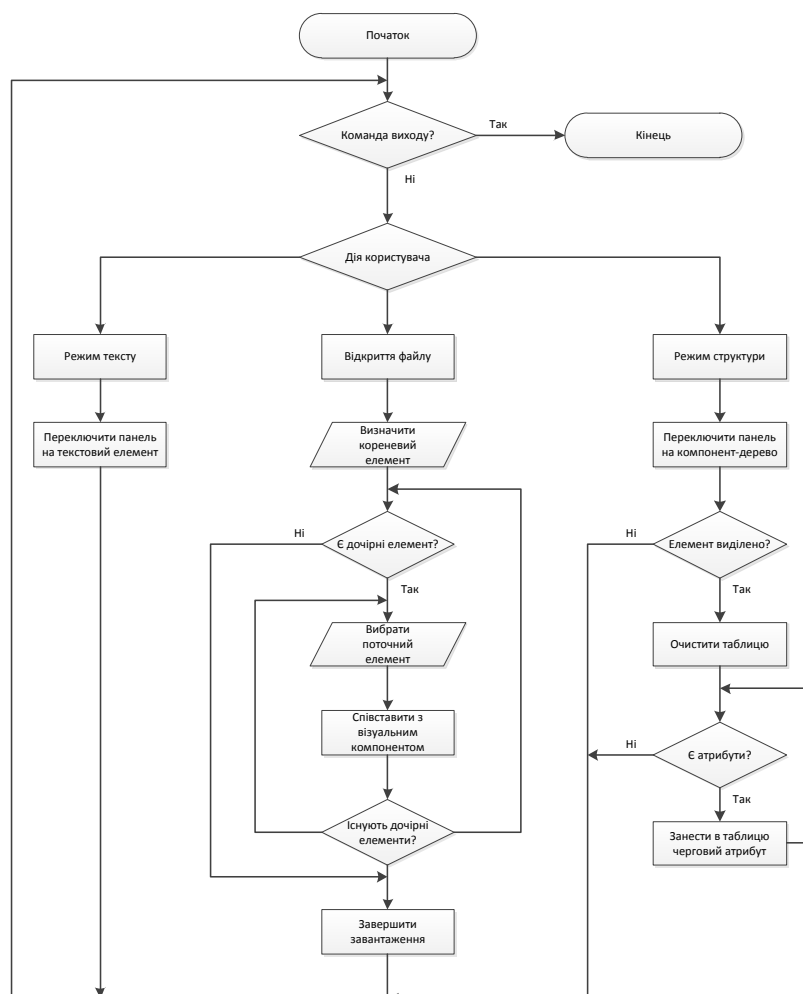


Рисунок 3 – Блок-схема алгоритму структурованої візуалізації XML-файлів

## Висновки

У даній роботі розроблено алгоритм структурованої візуалізації XML-файлів. Також, проведено аналіз XML-документа та виконано перевірку його на правильність, описано виконання рекурсивного завантаження всіх елементів документа, що аналізується.

## Список використаних джерел:

- 1.Трояновська Т. І. Розробка протоколу за стандартом SCORM для обміну даними між складовими СКП / Т. І. Трояновська // «Dynamika naukowych badan» : IX Міжнародна науково-практична конференція, 7–15 липня 2013 р. : тези доповідей. – Przemysl : «Nauka i studia», 2013. – С. 72–76. – ISBN 978–966–8736–05–6.
- 2.Гороховський О. І. Розробка ПЗ для автоматизації розробки SCORM-сумісних курсів / О. І. Гороховський, Т. І. Трояновська // «Стратегія якості у промисловості та освіті 2013» : Дев'ята міжнародна конференція, 31.05–07.06 2013 р. : тези доповідей. – Дніпропетровськ–Варна : «Економіка» – С. 404–407. – ISBN 978–966–02–5426–8.
- 3.M. Kay XSL Transformations 2.0 – W3C Working Draft, 2005.
- 4.E. R. Harold, W. Scott Means XML in a Nutshell – O'Reilly, 2005.
- 5.B. McLaughlin, J. Edelson Java and XML – O'Reilly, 2003.