**ФЕДЕРАЛЬНОЕ АГЕНТСТВО СВЯЗИ**

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ**

**«САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ТЕЛЕКОММУНИКАЦИЙ ИМ. ПРОФ. М.А. БОНЧ-БРУЕВИЧА»**

**(СПбГУТ)**

Факультет информационных систем и технологий

Кафедра информационных управляющих систем

**Дисциплина:** Корпоративные сетевые технологии

**Лабораторная работа №2**

Способы создания XML-документов. Проверка документа на правильность построения. Поиск ошибок в неправильно построенном документе

Ответственный исполнитель

Студент группы ИСТ–731

\_\_\_\_\_\_\_\_\_\_\_\_ Коханчик И.М.

«\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_\_\_ 2021 г.

Санкт-Петербург

2021

**Цель лабораторной работы:**

Целью работы является получение сведений о способах создания XML-документов, проверки на правильность их построения, типичных ошибках и их поиску.

**Задачи:**

- Изучить способы создания и проверки на правильность построения XML-документов;

- Найти и исправить ошибки в предоставленном документе;

- Дополнить документ в соответствии с правилами записи.

**1 Способы создания и проверки на правильность построения XML-документов**

Для создания документа можно использовать программу блокнот. При сохранении документа часто используют расширение xml. Необходимо учитывать тот факт, что документов, "написанных на XML", не может быть в принципе - в любом случае авторы документа для его разметки используют основанный на стандарте XML (т.н. XML-производный) язык, но не сам XML. Поэтому при сохранении созданного файла можно выбрать для него какое-то подходящее названию расширение.

Полученный файл необходимо проверить на правильность. Стандартом определены два уровня правильности документа XML:

• Правильно построенный (Well-formed). Правильно построенный документ соответствует всем общим правилам синтаксиса XML, применимым к любому XML-документу. И если, например, начальный тег не имеет соответствующего ему конечного тега, то это неправильно построенный документ XML. Документ, который неправильно построен, не может считаться документом XML; XML-процессор (парсер) не должен обрабатывать его обычным образом и обязан классифицировать ситуацию как фатальная ошибка.

• Действительный (Valid). Действительный документ дополнительно соответствует некоторым семантическим правилам. Это более строгая дополнительная проверка корректности документа на соответствие заранее определённым, но уже внешним правилам, в целях минимизации количества ошибок, например, структуры и состава данного, конкретного документа или семейства документов. Эти правила могут быть разработаны как самим пользователем, так и сторонними разработчиками, например, разработчиками словарей или стандартов обмена данными. Обычно такие правила хранятся в специальных файлах — схемах, где самым подробным образом описана структура документа, все допустимые названия элементов, атрибутов и многое другое. И если документ, например, содержит не определённое заранее в схемах название элемента, то XML-документ считается недействительным; проверяющий XML-процессор (валидатор) при проверке на соответствие правилам и схемам обязан (по выбору пользователя) сообщить об ошибке.

**2 Нахождение и исправление ошибок**

В XML-документе записан следующий код:

<?XML version = 1.0 encoding = WINDOWS-1251?>

<DVD жанр = боевик>

<название>Гладиатор</название>

<год знач = "2000">

<страна>США, Великобритания</страна>

<режиссер>Ридли Скотт</Режиссер>

<DVD жанр = фантастика>

<название>Зеленая миля</название>

<год знач = "1999">

<Страна>США</страна>

<режиссер>Фрэнк Дарабонт</режиссер>

</DVD>

Поиск ошибок будет произведен с помощью браузера Mozilla Firefox. При первой проверке обнаружена следующая ошибка, представленная на рисунке 1.

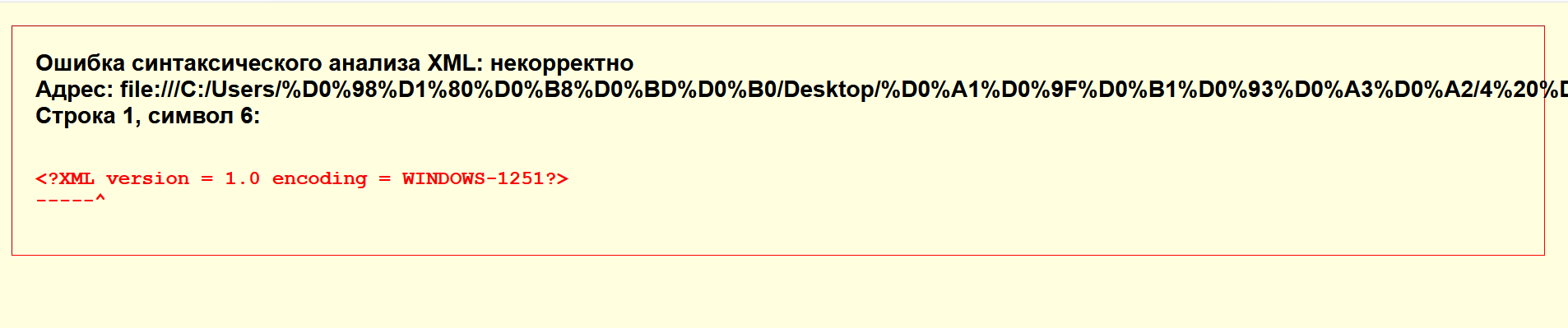


Рисунок 1

При этом способе проверки обнаруживается по одной ошибке за раз. Пример еще одной ошибки представлен на рисунке 2.

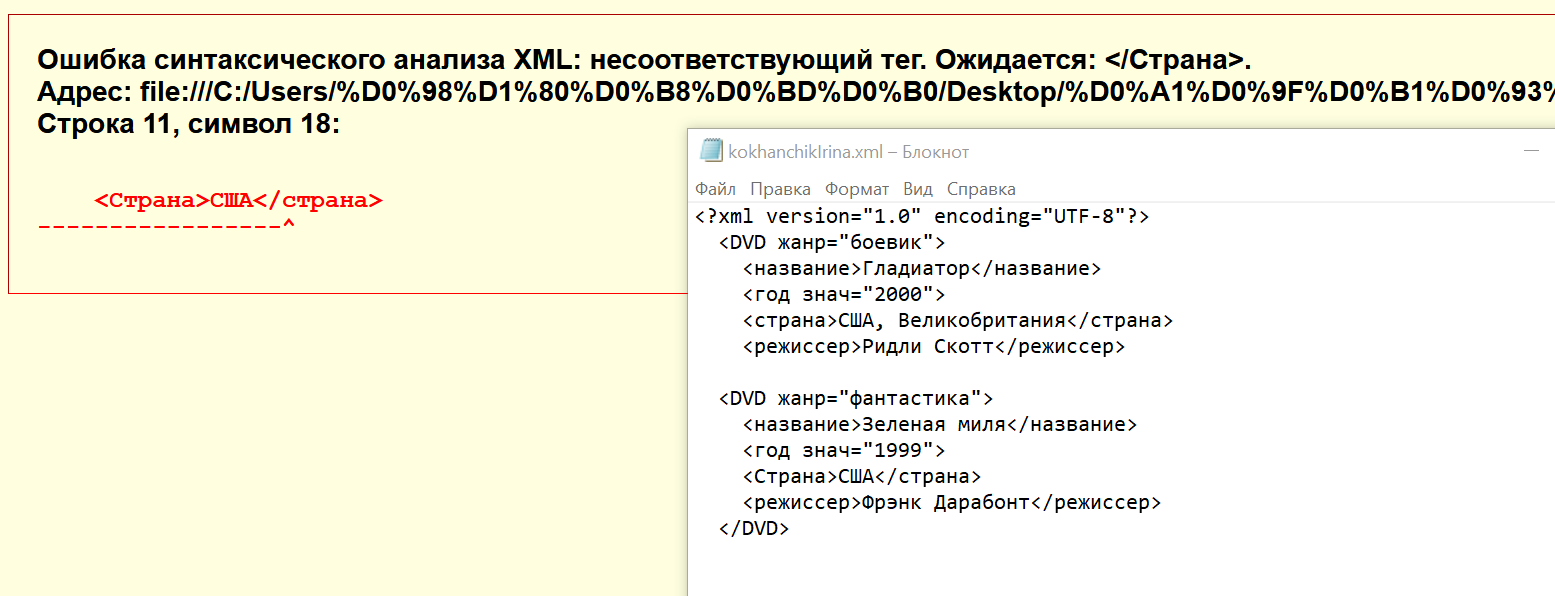


Рисунок 2

В ходе работы были обнаружены все ошибки. Проверка и исправленный код представлен на рисунке 3.

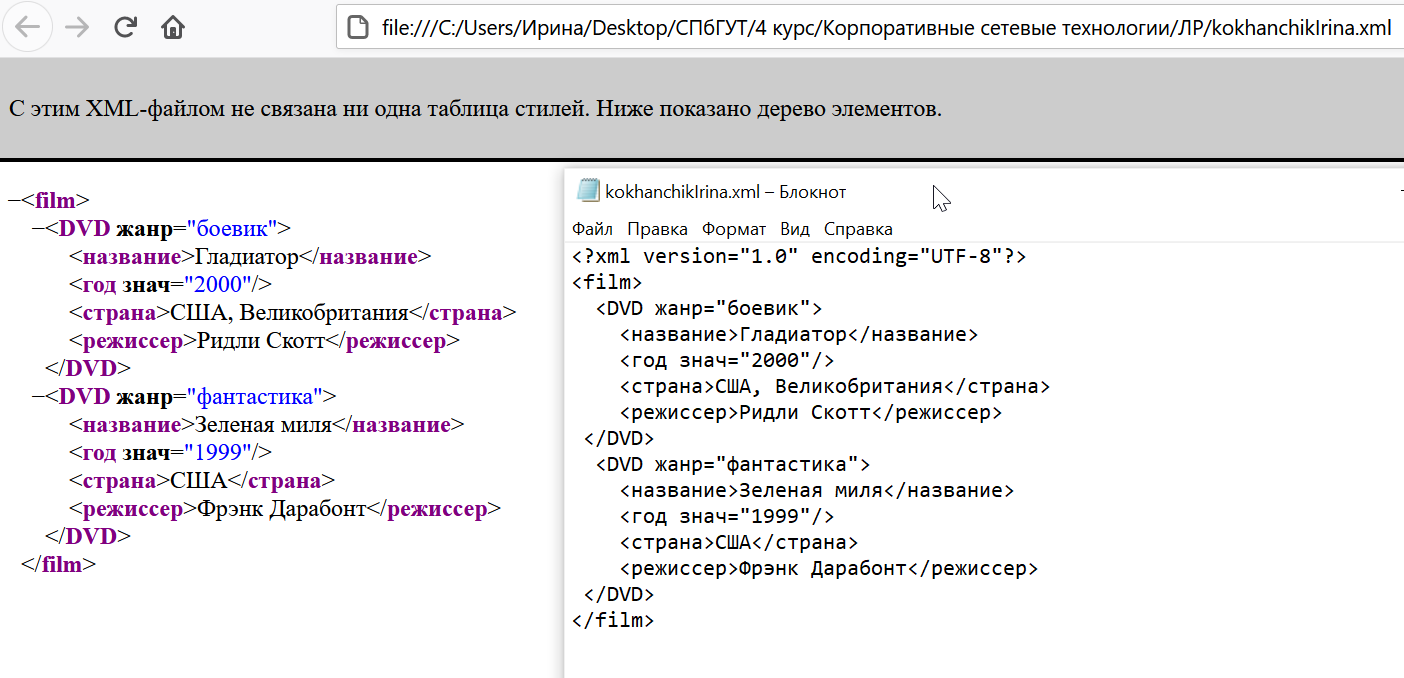


Рисунок 3

**3 Дополнение документа**

В код документа были добавлены фильмы как DVD, так и BD по заданию.

Код документа представлен ниже:

<?xml version="1.0" encoding="UTF-8"?>

<film>

<DVD жанр="боевик">

<название>Гладиатор</название>

<год знач="2000"/>

<страна>США, Великобритания</страна>

<режиссер>Ридли Скотт</режиссер>

</DVD>

<DVD жанр="фантастика">

<название>Зеленая миля</название>

<год знач="1999"/>

<страна>США</страна>

<режиссер>Фрэнк Дарабонт</режиссер>

</DVD>

<BD жанр="комедия">

<название>Зависнуть в Палм-Спрингс</название>

<год знач="2020"/>

<страна>США</страна>

<режиссер>Макс Барбаков</режиссер>

</BD>

<BD жанр="триллер">

<название>Суд над чикагской семеркой</название>

<год знач="2020"/>

<страна>США</страна>

<режиссер>Аарон Соркин</режиссер>

</BD>

<DVD жанр="фантастика">

<название>Довод</название>

<год знач="2020"/>

<страна>США</страна>

<режиссер>Кристофер Нолан</режиссер>

</DVD>

<DVD жанр="драма">

<название>Маленькие женщины</название>

<год знач="2019"/>

<страна>США</страна>

<режиссер>Грета Гервиг</режиссер>

</DVD>

<DVD жанр="криминал">

<название>Джентльмены</название>

<год знач="2019"/>

<страна>США</страна>

<режиссер>Гай Ричи</режиссер>

</DVD>

<DVD жанр="биография">

<название>Манк</название>

<год знач="2020"/>

<страна>США</страна>

<режиссер>Девид Финчер</режиссер>

</DVD>

<DVD жанр="военный">

<название>1917</название>

<год знач="2019"/>

<страна>США</страна>

<режиссер>Сэм Мендес</режиссер>

</DVD>

<DVD жанр="драма">

<название>Еще по одной</название>

<год знач="2020"/>

<страна>Дания</страна>

<режиссер>Томас Винтерберг</режиссер>

</DVD>

<DVD жанр="ужасы">

<название>Маяк</название>

<год знач="2019"/>

<страна>США</страна>

<режиссер>Роберт Эггерс</режиссер>

</DVD>

</film>

Проверка синтаксиса представлена на рисунке 4.

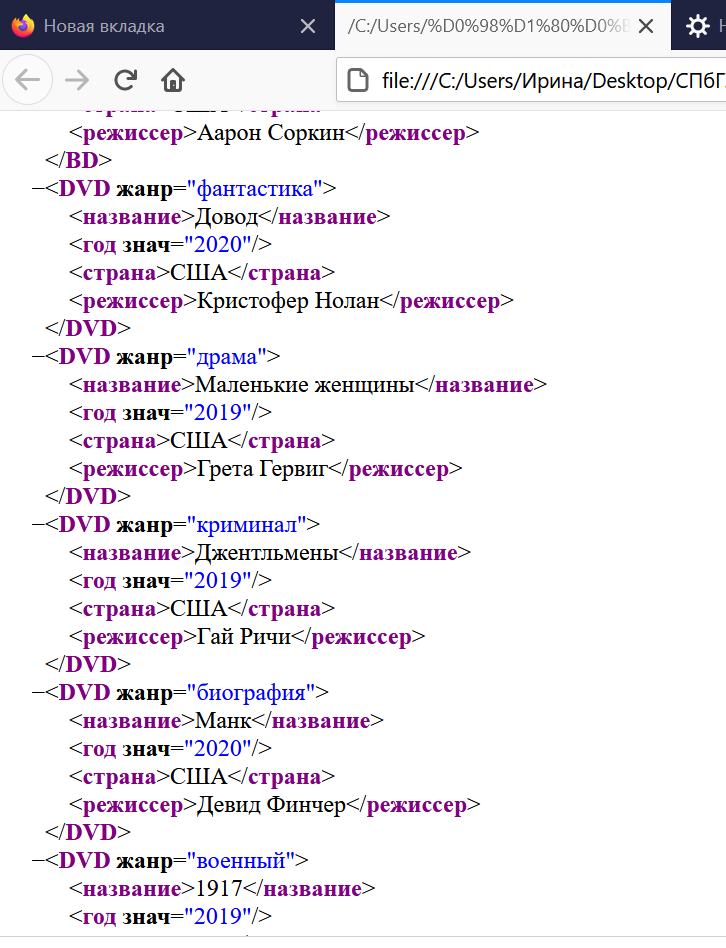


Рисунок 4