

Установка и использование libxml2

In progress...

Установка libxml2

Вариант 1

Используя стандартный репозиторий.
В Fedora:

```
#sudo yum install libxml2
```

Вариант 2

Из исходников:

```
# wget ftp://xmlsoft.org/libxml2/libxml2-2.6.30.tar.gz
# tar -zxvf libxml2-2.6.30.tar.gz
# cd libxml2-2.6.30/
# ./configure && make && make install
# cd /usr/include
# ln -s libxml2/libxml libxml
```

Основные функции

Прототипы основных функций, которые необходимы для обхода XML, содержатся в заголовочных файлах **parser.h** и **tree.h**.

Для проверки версии XML используется **LIBXML_TEST_VERSION**

Для сопоставления файла с XML используется функция **xmlReadFile** (первый параметр функции - имя файла для разбора, два другие не знаю), получается так:

```
xmlDocPtr doc = xmlReadFile(argv[1], NULL, 0);
```

Переход между элементами

Дальше следует получить корень документа -

```
xmlNodePtr root = xmlDocGetRootElement(doc);
```

Переход к следующему узлу (если узел не соержжит дочерних):

```
xmlNodePtr cur_node = root;  
cur_node = cur_node->next;
```

Переход к дочернему узлу:

```
cur_node = cur_node->children;
```

Переход к отцу:

```
cur_node = cur_node->parent
```

Получение данных тега

У тега есть имя, атрибуты и содержание (информация). Чтобы получить их используются следующие конструкции ¹⁾:

Получение имени тега:

```
printf( "%s", cur_node->name);
```

Работа с атрибутами (получение названий и значений и переход между атрибутами)

```
xmlAttrPtr attr = cur_node->properties; // берем атрибут  
while(attr)  
{  
    printf(" %s", attr->children->name); печать имени атрибута  
    printf(" %s", attr->children->content); печать значения атрибута  
    attr = attr->next;//на следующий атрибут  
}
```

Получение содержания тега:

```
printf(" %s",xmlNodeGetContent(cur_node));
```

Пример использования

Программа сделает обход вглубину всех тегов XML-файла

```
#include<stdio.h>  
#include<string.h>  
#include<libxml/parser.h>  
#include<libxml/tree.h>  
  
xmlDocPtr doc = NULL;//указатель на документ
```

```
void rec(xmlNodePtr node)
{
    xmlNodePtr cur_node = node;
    while (cur_node)
    {
        if (cur_node->type == XML_ELEMENT_NODE) //проверяем что попали на узел
        {
            printf( "node: %s\n", cur_node->name);
        }
        rec(cur_node->children);
        cur_node = cur_node->next;
    }
}

int main(int argc, char **argv)
{
    LIBXML_TEST_VERSION;
    doc = xmlReadFile(argv[1], NULL, 0);
    if(doc == NULL)
    {
        printf("Error parsing file %s\n", argv[1]);
        exit(1);
    }
    xmlNodePtr root = xmlDocGetRootElement(doc);
    rec(root);
    xmlFreeDoc(doc);
    return 0;
}
```

Команда компиляции:

```
gcc -lxml2 -o res main.c
```

Пример файла для разбора:

```
<?xml version="1.0"?>
<gjob:Helping xmlns:gjob="http://www.gnome.org/some-location">
  <gjob:Jobs>

    <gjob:Job>
      <gjob:Project ID="3"/>
      <gjob:Application>GBackup</gjob:Application>
      <gjob:Category>Development</gjob:Category>

      <gjob:Update>
        <gjob:Status>Open</gjob:Status>
        <gjob:Modified>Mon, 07 Jun 1999 20:27:45 -0400 MET
DST</gjob:Modified>
        <gjob:Salary>USD 0.00</gjob:Salary>
      </gjob:Update>
```

```
<gjob:Developers>
  <gjob:Developer>
  </gjob:Developer>
</gjob:Developers>

<gjob>Contact>
  <gjob:Person>Nathan Clemons</gjob:Person>
  <gjob:Email>nathan@windsofstorm.net</gjob:Email>
  <gjob:Company>
  </gjob:Company>
  <gjob:Organisation>
  </gjob:Organisation>
  <gjob:Webpage>
  </gjob:Webpage>
  <gjob:Snailmail>
  </gjob:Snailmail>
  <gjob:Phone>
  </gjob:Phone>
</gjob>Contact>

</gjob:Job>

</gjob:Jobs>
</gjob:Helping>
```

Запуск и результат:

```
# ./res exsample.xml
node: Helping
node: Jobs
node: Job
node: Project
node: Application
node: Category
node: Update
node: Status
node: Modified
node: Salary
node: Developers
node: Developer
node: Contact
node: Person
node: Email
node: Company
node: Organisation
node: Webpage
node: Snailmail
node: Phone
```

для всех текстов справедливо `xmlNodePtr cur_node;`

From:

<http://wiki.osll.ru/> - **Open Source & Linux Lab**

Permanent link:

http://wiki.osll.ru/doku.php/howtos:libxml2_install_and_use

Last update: **2016/08/07 01:21**

