



Feladat

1. A tárgy honlapjáról töltsük le a `web.docbook` állományt. Az állomány tartalma a következő:

```
<article>
<title>About the Web</title>

<para>
This is an article about the World Wide Web.
The World Wide Web is a collection of documents that are linked to
one another. The Web is <emphasis>not</emphasis> the same as the
Internet. The Internet is a world-wide network of networks, and it
does far more than simply serve up Web pages.
</para>

<para>Tim Berners-Lee, the inventor of the World Wide Web, put special
emphasis on the portability of web pages. Rather than create a
proprietary format, he made Web pages dependent only upon plain ASCII
text.</para>

<para>
Web pages are written in a markup language called HTML. Here is what it
looks like. The &lt; and &gt; mark off elements.
</para>

<listing>
<lt;body>&gt;
<lt;div id="top-navig"&gt;
<lt;a id="top"&gt;&lt;/a>&gt;
<lt;a href="index.html"&gt;CIT 040 Index&lt;/a>&gt;
&amp;&gt;
Assignment 1
<lt;/div>&gt;

<lt;h1>Assignment 1&lt;/h1>&gt;
<lt;p>&gt;This exercise shows you how to use the two computer environments that
you will use in this class. You will:&lt;/p>&gt;
<lt;ol class="upper-roman"&gt;
<lt;li>&gt;Set up your directories on Windows. This is
where you will write your HTML documents.&lt;/li>&gt;
<lt;/ol>&gt;
</listing>

<para>It looks difficult, but it is possible to learn HTML in a few
weeks. <emphasis>You, too can create web pages for viewing by
friends and family!</emphasis>
Note that, in our listing, we had to encode &gt; as &amp;gt;.
</para>
</article>
```

Ez egy Docbook állomány, felépítése hasonlít a weben használt HTML-hez. Könyvek, dokumentumok írásához használt formátum.

Egy Docbook állomány valamilyen szövegből áll, ahol a szöveg egyes részei „tag”-ek között szerepel(het)nek. Egy nyitó tag egy csúcsos zárójelek között szereplő parancs, pl. `<para>` (jelentése: paragrafus kezdete). Az ehhez tartozó záró tag: `</para>` (jelentése: paragrafus vége).

A továbbiakban ezt a szöveget fogjuk módosítani a `sed` segítségével. Ehhez írjunk egy `sed` parancsfájlt. A parancsfájlnak tetszőleges Docbook állományon is működni kell, vagyis nem számíthatunk arra, hogy egy bizonyos tag csak egy bizonyos sorban fordul elő.

2. Az `<article>` és `</article>` sorokat töröljük.
3. Cseréljük le a `<title>` tag-et `Title: -re`, a `</title>` tag-et pedig üres sztringre.
4. A `<para>` és `</para>` tag-eket cseréljük le üres sztringre. Ha az eredményül kapott sor üres, töröljük a sort. (Használjuk a kacsos zárójeleket több parancs összefogására.)

5. Az `<emphasis>` és `</emphasis>` tag-eket cseréljük le `*` -ra. Például:

```
This is a <emphasis>great</emphasis> bargain.
```

Átalakítás után így fog kinézni:

```
This is a *great* bargain.
```

6. A `web` szót mindenhol cseréljük le `Web` -re.
7. A `<listing>` tag-eket cseréljük le `---begin listing` -re.
8. A `</listing>` tag-eket cseréljük le `---end listing` -re.
9. A `<listing>` és `</listing>` sorok között a következő műveleteket hajtsuk végre (ehhez a parancsokat kapcsos zárójelek között kell összefogni):
- Az `<` összes előfordulását cseréljük le `<` -re.
 - Az `>` összes előfordulását cseréljük le `>` -re.
 - Az `&` összes előfordulását cseréljük le `&` -re.

Megjegyzés #1: a cseréket ebben a sorrendben hajtsuk végre, különben rossz eredményt kapunk!

Megjegyzés #2: ha az `&` jelet a csere jobb oldalán akarjuk használni (pl. az `and` -et `&` -re akarjuk cserélni), akkor ezt a következőképpen tehetjük meg: `s/and/\&/`.

10. A végeredménynek így kell kinéznie (`web_eredmeny.txt` a honlapon):

```
Title: About the Web

This is an article about the World Wide Web.
The World Wide Web is a collection of documents that are linked to
one another. The Web is *not* the same as the
Internet. The Internet is a world-wide network of networks, and it
does far more than simply serve up Web pages.

Tim Berners-Lee, the inventor of the World Wide Web, put special
emphasis on the portability of Web pages. Rather than create a
proprietary format, he made Web pages dependent only upon plain ASCII
text.

Web pages are written in a markup language called HTML. Here is what it
looks like. The &lt; and &gt; mark off elements.

---begin listing
<body>
<div id="top-navig">
<a id="top"></a>
<a href="index.html">CIT 040 Index</a>
&gt;
Assignment 1
</div>

<h1>Assignment 1</h1>
<p>This exercise shows you how to use the two computer environments that
you will use in this class. You will:</p>
<ol class="upper-roman">
<li>Set up your directories on Windows. This is
where you will write your HTML documents.</li>
</ol>
---end listing

It looks difficult, but it is possible to learn HTML in a few
weeks. *You, too can create Web pages for viewing by
friends and family!*
Note that, in our listing, we had to encode &gt; as &amp;gt;.
```

11. Ha készen vagyunk, tömörítsük be mindhárom fájlt egy `.zip` állományba, majd küldjük el a gyakorlatvezető email címére. Tüntessük fel a nevünket, ill. azt is, hogy melyik csoportba járunk.