

A webböngészők a világhálón található oldalakat jelenítik meg számítógépeinken. Ehhez minden HTML oldal tartalmának értelmezésekor egy dokumentumfát készítenek. A dokumentumfa gyökere a <HTML> elem, minden további elem pedig az őt tartalmazó elem leszármazottja. Példaként tekintsük az alábbi HTML forrást, és a hozzá tartozó dokumentumfát:

```

<html>
<head>
<title>Én kis oldalam</title>
</head>
<body>
<h3>Üdvözöllek látogató!</H3>
<table class="oszt">
<tr>
<td></td>
<td><a href="#">
Ide <span id="ide">katt</span>!
</a></td>
</tr>
</table>
<h4>Látogass ide máskor is...</h4>
</body>
</html>

```

```

html
|
+--head
| |
| +--title
|
+--body
|
+--h3
|
+--table
| |
| +--tr
| | |
| | +--td
| | | |
| | | +--img
| | |
| | +--td
| | |
| | | +--a
| | | |
| | | | +--span
|
+--h4

```

Írjunk programot, amely egy HTML 4.01 Strict DTD verziójú (tehát érvénytelenített formázásoktól mentes és kereteket nem tartalmazó), hibátlan oldal dokumentumfáját elkészíti. Az érvényes HTML elemek a

<http://www.w3.org/TR/html401/index/elements.html>

oldalon találhatóak (csak a D és L jelzés nélküliek, összesen 80 elem). A program olvassa be a parancssor első argumentumaként megadott nevű forrásállományból az oldalt tartalmazó szöveget, majd írja ki a parancssor második argumentumaként megadott nevű kimeneti szöveges állományba a fenti formában a dokumentumfát.

Beküldendő a program forráskódja (s38.pas, s38.cpp, ...), valamint a program rövid dokumentációja (s38.txt, s38.pdf, ...), amely tartalmazza a megoldás rövid leírását, és megadja, hogy a forrásállomány melyik fejlesztő környezetben fordítható.