

Transformácia jednoduchého značkovacieho jazyka do HTML (skúšobné zadanie OS ZS 2021/22)

November 14, 2023

1 Zadanie

Rozšírte `markdown.sh` z domáceho zadania tak, aby podporoval tieto ďalšie syntaktické konštrukty.

1.1 URL

Kdekoľvek na riadku sa nachádza retazec typu `<https://text>`, nahradte ho podľa pravidla

`<https://text>` → `https://text`

napríklad riadok

```
Google: <https://google.sk/> xxx
```

sa transformuje na

```
Google: <a href="https://google.sk/">https://google.sk/</a> xxx
```

1.2 Zoznamy

Začnite s tým, že zmeníte

```
while read LINE
do
```

na

```
while IFS= read LINE
do
```

Je to kvôli tomu, že príkaz `read LINE` za normálnych okolností zahadzuje počiatočné medzery na riadku.

Transformujte každú sekvenciu riadkov typu

```
Niečo pred
- hlava
- trup
- nohy
Niečo po
```

na sekvenciu typu

```
Niečo pred
<ol>
<li>hlava</li>
<li>trup</li>
<li>nohy</li>
</ol>
Niečo po
```

To znamená

- Ak narazíte na prvý výskyt \square ¹ na začiatku riadku vypíšte najprv riadok ``
- Riadky v zozname transformujte podľa pravidla

¹ \square znamená medzeru.

`□-□text → text`

- Ak narazíte na nejaký riadok ktorý nezačína `□-□` a predtým bol zoznam, je to koniec zoznamu a pred ďalším spracovaním vypíšete ``, aby ste ten zoznam správne ukončili.

2 Pomôcky

2.1 Použite spätné referencie

V prvej časti je rozumné použiť `sed /regex/replacement/`, pričom v *replacement* budete referencovať cez `\1` časť *regex*, ktorú ste vyznačili pomocou `\(...\)`.

2.2 Pomocná premenná

Pre spracovanie zoznamov potrebujete nejakú pomocnú premennú, kde si budete pamätať, či ste uprostred zoznamu alebo nie, povedzme `ZOZ`. Jej rovnosť nejakej hodnoty môžete zistiť pomocou `if` a `test`

```
if test "$ZOZ" = "y"
then
```

prípadne (pozor na medzery!)

```
if [ "$ZOZ" = "y" ]
then
```

2.3 Vstup (příklad)

0 úpadku doby

Karel Čapek

Před jeskyní bylo ticho. Mužové odešli hned zrána mávajíce oštěpy směrem k __Blansku__ nebo k __Rájci__, kde bylo vystopováno stádo sobů; ženy zatím po lese sbíraly bobule klikvy a jen časem bylo slyšet jejich ječivý pokřik a repetění; děti se nejspíš čabraly dole v potoce - ostatně kdo by se těch harantů dohlídal, holoty darebné a zvlčilé. A tož starý pračlověk _Janeček_ klímal v tom vzácném tichu na mírném slunci říjnovém; po pravdě řečeno chrápal a v nose mu hvízdalo, ale dělal, jako by nespál, nýbrž jako by střežil jeskyni kmene a panoval nad ní, jakož se sluší na starého náčelníka.

Janečková rozložila čerstvou kůži _medvědí_ a jala se ji oškrabovat ostrým pazourkem. To se musí dělat důkladně, píd' po pídi - a ne jako to dělá mladá, napadlo starou Janečkovou; to famfárum to jen tak halabala odrbe a už zas běží muchlat a mukat se s dětmi - takové kůže, myslí si stará Janečková, nic nevydrží, kdepak, ztyří nebo se zapaří; ale já se jí do ničeho plést nebudu, myslí si paní Janečková, když jí to neřekne syn - Jen co je pravda, šetřit mladá neumí. A tady je ta kožišina propíchnuta, zrovna uprostřed hřbetu! Lidi drazí, trne stará paní, kterýpak nekola toho medvěda píchal do zad? Vždyť se tím zkazí celá kůže! To by můj jakživ neudělal, říká si stařena roztrpčeně, ten se vždycky strefoval do krku -

Stránka predmetu OS <<https://www.math.sk/wiki/OperacneSystemy>> je hento.

Dve URL na riadku <<https://www.google.com/>> aj toto <<https://www.stuba.sk>> je URL.

Zoznam nasleduje

- Zoznam
- nejakých
- vecí

a toto je nejaký iný text.

2.4 Výstup (příklad)

```
<!DOCTYPE html>
<html>
<head>
  <meta http-equiv="Content-type" content="text/html; charset=UTF-8" />
</head>
<body>
<h1>0 úpadku doby</h1>
<h2>Karel Čapek</h2>
Před jeskyní bylo ticho. Mužové odešli hned zrána mávajíce oštěpy směrem k
<strong>Blansku</strong> nebo k <strong>Rájci</strong>, kde bylo vystopováno stádo sobů; Ženy zatím po lese
sbíraly bobule klikvy a jen časem bylo slyšet jejich ječivý pokřik a repetění;
dětí se nejspíš čabraly dole v potoce - ostatně kdo by se těch harantů
dohlídal, holoty darebné a zvlčilé. A tož starý pračlověk <em>Janeček</em> klímal v tom
vzácném tichu na mírném slunci Fijnovém; po pravdě řečeno chrápal a v nose mu
hvízdalo, ale dělal, jako by nespál, nýbrž jako by střežil jeskyni kmene a
panoval nad ní, jakož se sluší na starého náčelníka.
<p>
<em>Janečková</em> rozložila čerstvou kůži <em>medvědí</em> a jala se ji oškrabovat ostrým
pazourkem. To se musí dělat důkladně, píd po pidi - a ne jako to dělá mladá,
napadlo starou Janečkovou; to famfárum to jen tak halabala odrbe a už zas běží
muchlat a muckat se s dětmi - takové kůže, myslí si stará Janečková, nic
nevydrží, kdepak, ztyří nebo se zapafí; ale já se jí do ničeho plést nebudu,
myslí si paní Janečková, když jí to neřekne syn - Jen co je pravda, šetřit
mladá neumí. A tady je ta kožišina propíchnuta, zrovna uprostřed hřbetu! Lidi
drazí, trne stará paní, kterýpak nekoľa toho medvěda píchal do zad? Vždyť se
tím zkazí celá kůže! To by můj jakživ neudělal, říká si stařena roztrpčeně, ten
se vždycky střefoval do krku -
<p>
Stránka predmetu OS <a href="https://www.math.sk/wiki/OperacneSystemy">https://www.math.sk/wiki/OperacneSystemy</a> je hento.
<p>
Dve URL na riadku <a href="https://www.google.com/">https://www.google.com/</a> aj toto <a href="https://www.stuba.sk">https://www.stuba.sk</a> je URL.
<p>
Zoznam nasleduje
<ul>
<li>Zoznam</li>
<li>nejakých</li>
<li>vecí</li>
</ul>
a toto je nejaký iný text.
<p>
</body>
</html>
```