



Unix/Linux

ISEN - AP4

Y - Projet - My Horrible XML parser

2019 - J. Hochart

Objectif

- XML de référence:
https://www.w3schools.com/xml/plant_catalog.xml
- Créer un parser spécialisé pour la conversion de ce schéma en CSV
- Toujours un saut de ligne entre les balises
- Le script n'aura qu'à traiter le XML de référence, ou une sous-partie (à l'exception des balises optionnelles)

Specs

```
root@debian:/home/jho# head -12 plant_catalog.xml
```

```
<?xml version="1.0" encoding="UTF-8"?>
<CATALOG>
  <PLANT>
    <COMMON>Bloodroot</COMMON>
    <BOTANICAL>Sanguinaria canadensis</BOTANICAL>
    <ZONE>4</ZONE>
    <LIGHT>Mostly Shady</LIGHT>
    <PRICE>$2.44</PRICE>
    <AVAILABILITY>031599</AVAILABILITY>
  </PLANT>
  <PLANT>
```

Specs

```
root@debian:/home/jho# cat plant_catalog.xml | mhxmlp.sh
```

```
Bloodroot;Sanguinaria canadensis;4; Mostly Shady;$2.44;031599
```

```
Columbine;Aquilegia canadensis;3;Mostly Shady;$9.37;030699
```

```
...
```

Cas particuliers

- Il est possible que, pour une plante, un ou plusieurs champs soient absents dans le XML (pas de tag)
- Dans ce cas, la valeur de la colonne dans le csv est la chaîne vide
- Seul le champ COMMON est obligatoire et sera toujours présent dans le XML

Outils imposés

- Shell builtins (while, read, do, echo, ...)
 - Cat
 - Grep
 - Cut
 - Sed
-
- Rien d'autre!

Arch

Build	Debian 9.8 Bash 4.4 Coreutils 8
Run	= Build

Interface

Call	<code>./mhxmsp.sh</code>
Arguments	Aucun. Le script lit sur stdin

Cas particuliers

Cas	Return
Mauvais nombre d'arguments	5
Rien sur stdin	11
Parse error (le xml en entrée est invalide)	19
Succès	0

Livraison

Mode	1 livraison par étudiant
Livrable	1 fichier source nommé "mhxmlp.txt" (renommer l'extension sh en txt pour l'envoi par mail)
Transmission	Par email (jul@hochart.fr), avec le fichier en pièce jointe non compressé Titre: mhxmlp prénom nom
Limite (envoi)	04/05/2019 @ 20:00 (Lille)

Livraison

Étapes de correction	<pre>\$ cp /rendu/\$NAME/mhxmlp ./ \$ cat sample.xml mhxmlp.sh</pre>
Notation	<p>Default: 0 Livraison conforme: 4 Pass test 1: 10 Pass test 2: 12 Pass test 3: 14 Pass test 4: 16 Pass test 5: 18 Pass test 6: 20</p>