

Specifica Progetto Appelli Febbraio 2019

Descrizione

Sviluppare una applicazione di sistema Unix/Linux chiamata `swordx` che sia in grado di leggere un insieme di file (di testo) da una o più sorgenti e che produca in output un file di testo contenente la lista delle parole (ordinate in ordine alfabetico) che occorrono nei file letti con la relativa occorrenza. Per parola si intende ogni sequenza costituita da soli caratteri alfanumerici (caratteri o lettere). Le parole non sono sensibili al maiuscolo.

La sinossi del programma è:

```
swordx [options] [inputs]
```

Dove `[inputs]` è una sequenza di file e/o directory, mentre `[options]` sono una sequenza di parametri riportati sotto.

Dettagli

Feature 1

Il programma è in grado di trattare invocazioni del tipo:

```
swordx path/file
```

Il risultato dell'elaborazione è salvata nel file `swordx.out` salvato nella directory corrent.

Esempio: Ricevuto in input il file `test.txt` (https://drive.google.com/open?id=1W10ctaVDIzQ-CPODa_P65TZQgQiqAFQJ) viene prodotto il seguente file `swordx.out` (https://drive.google.com/open?id=1vJrNWczm2DusLJ1rFH_F0xVcEIBWi1S).

Feature 2

Come input può essere passata una directory:

```
swordx path/directory
```

In questo caso il file `swordx.out` conterrà le statistiche ottenute elaborando tutti i file (regolari) contenuti nella directory. Nell'analisi, quindi, vengono esclusi i link e le directory.

Feature 3

E' possibile richiamare `swordx` passando più input:

```
swordx <input1> <input2> ... <inputn>
```

Viene generato un singolo file `swordx.out` contenente le statistiche collezionate.

Feature 4

E' possibile indicare il nome del file dove devono essere salvate le statistiche collezionate:

```
swordx -output <file> [inputs] swordx -o <file> [inputs]
```

Le statistiche collezionate vengono salvate nel file `<file>`.



Feature 5

E' possibile passare a `swordx` i seguenti parametri aggiuntivi:

1. `-help` e `-h` : stampa a video l'help del programma;
2. `-recursive` e `-r` : se viene passato questo argomento, nella collezione delle informazioni di una directory vengono seguite tutte le sottodirectory;
3. `-follow` e `-f` : nell'elaborazione di una directory vengono seguiti i link;
4. `-exclude <file>` e `-e <file>` : `<file>` non viene considerato nell'elaborazione (questo parametro ha senso solo se si sta elaborando una directory);
5. `-alpha` e `-a` : vengono considerate nella statistica solo le parole contenenti caratteri alfabetici;
6. `-min <num>` e `-m <num>` : vengono considerate nella statistiche solo le parole con una lunghezza maggiore o uguale a `<num>` ;
7. `-ignore <file>` e `-i <file>` : `<file>` è un elenco di parole (una per riga) che vengono ignorate nella statistica;
8. `-sortbyoccurrency` e `-s` : le parole vengono inserite nel file di output ordinate per numero di occorrenze.
9. `-log <file>` e `-l <file>` : genera un file di log (identificato da `<file>`) dove viene riportate viene riportata, per ogni file analizzato viene riportata una riga della forma `<name> cw iw time` dove
 - a. `<name>` è il nome del file analizzato;
 - b. `cw` è il numero delle parole registrate;
 - c. `iw` è il numero delle parole ignorate;
 - d. `time` è il tempo necessario all'elaborazione del file.
10. `-update <file>` se presente il file `swordx.out` (o il file indicato con l'opzione `-o`) scrive nel file `<file>` , con lo stesso formato del file di output, l'elenco delle parole con le rispettive variazioni rispetto all'esecuzione precedente del programma.

Feature 6

Consentire di specificare valori parametri di input usando espressioni regolari.

 [didattica/triennale/so/ay_1718/progettofebbraio.txt](#)  Last modified: 2018/11/07 17:19 by michele

Except where otherwise noted, content on this wiki is licensed under the following license: CC Attribution-Share Alike 4.0 International (<http://creativecommons.org/licenses/by-sa/4.0/>)