Piuparts Broken Symlink Analysis

The Debian Piuparts program performs installation, upgrade, and removal tests against packages in the repositories. The piuparts.debian.org website provides a resource for collecting and reporting results of these tests.

The sid "successfully tested" list has a "but logfile contains broken symlinks" sublist. As of this writing (Nov 2011) 30,678 of the 31,352 'successfully tested' packages in sid report at least one broken symlink, flagged as a warning.

This script will collect broken symlink messages from Piuparts log files, and report the broken symlinks found, with a count of packages reporting each symlink.

This is a typical output, based on 30623 log files containing broken symlinks, collected 11/13/11:

```
# piubsl.py linkreport db
count broken symlink
29706 /etc/motd -> /var/run/motd
 6382 /etc/fonts/conf.d/30-defoma.conf -> /var/lib/defoma/fontconfig.d/fonts.conf
  877 /usr/lib/python2.6/dist-packages/python-support.pth -> ../../pymodules/python2.6/.path
  874 /usr/lib/python2.7/dist-packages/python-support.pth -> ../../pymodules/python2.7/.path
  757 /usr/share/pyshared/numpy/core/include/numpy/nummacro.h -> ../../../numarray/numpy/nummacro.h
  757 /usr/share/pyshared/numpy/core/include/numpy/ieeespecial.h -> ../../../numarray/numpy/ieeespecial.h
  757 /usr/share/pyshared/numpy/core/include/numpy/numcomplex.h -> ../../../numarray/numpy/numcomplex.h
  757 /usr/share/pyshared/numpy/core/include/numpy/libnumarray.h -> ../../../numarray/numpy/libnumarray.h
  757 /usr/share/pyshared/numpy/core/include/numpy/arraybase.h -> ../../../numarray/numpy/arraybase.h
  757 /usr/share/pyshared/numpy/core/include/numpy/cfunc.h -> ../../../numarray/numpy/cfunc.h
  536 /usr/lib/gcc/x86_64-linux-gnu/4.6/4.6 -> 4.6
  517 /usr/share/doc/libgtk-3-0/README.gz -> ../libgtk-3-common/README.gz
  399 /usr/share/dict/words -> /etc/dictionaries-common/words
  393 /usr/lib/ispell/default.aff -> /etc/dictionaries-common/default.aff
  393 /usr/lib/ispell/default.hash -> /etc/dictionaries-common/default.hash
  368 /usr/share/man/man1/2to3.1.gz -> 2to3-2.6.1.gz
  362 /usr/include/python2.6/numpy -> ../../lib/pymodules/python2.6/numpy/core/include/numpy
  362 /usr/include/python2.6_d/numpy -> ../../lib/pymodules/python2.6/numpy/core/include/numpy
  335 /usr/include/python2.7_d/numpy -> ../../lib/pymodules/python2.7/numpy/core/include/numpy
  335 /usr/include/python2.7/numpy -> ../../lib/pymodules/python2.7/numpy/core/include/numpy
  128 /usr/lib/xulrunner-1.9.1/dictionaries -> ../../share/hunspell
  107 /usr/share/doc/libgtk-3-bin/README.gz -> ../libgtk-3-common/README.gz
   91 /usr/share/doc/python-simplejson/_static/sidebar.js -> ../../../javascript/sphinxdoc/1.0/sidebar.js
...
```

The database used for this output, 'db', is included in the repository.

Usage

Main program help:

```
# piubsl.py -h
usage: piubsl.py [-h] [-v] {initdb,addlog,linkreport,linkdetail} ...

Analyze Piuparts broken symlinks

Using log files available under e.g.

    http://piuparts.debian.org/sid/broken_symlinks_issue.html

analyze the frequency and source of "Broken symlink" warnings

positional arguments
  {initdb,addlog,linkreport,linkdetail}
                        command help
    initdb              initialize the database file
```

```
    addlog                  add a log file to the database
    linkreport              print a summary of broken symlinks found
    linkdetail              print a summary of packages related to a broken link


optional arguments:
  -h, --help              show this help message and exit
  -v


use 'piubsl.py <cmd> -h for command-specific help
```

Initialize a database:

```
# piubsl.py initdb -h
usage: piubsl.py initdb [-h] dbfile


Initialize the sqlite database file to contain symlink stats


positional arguments:
  dbfile      the name of the database file


optional arguments:
  -h, --help  show this help message and exit
```

Add a log to the database:

```
# piubsl addlog -h
usage: piubsl.py addlog [-h] dbfile [logfiles [logfiles ...]]


parse a single Piuparts log file into the database


positional arguments:
  dbfile      the name of the database file
  logfiles    the name of the logfile (omit for stdin)


optional arguments:
  -h, --help  show this help message and exit
```

And output the link report:

```
# piubsl linkreport -h
usage: piubsl.py linkreport [-h] dbfile


Print a report of broken symlinks found in the log files


positional arguments:
  dbfile      the name of the database file


optional arguments:
  -h, --help  show this help message and exit
```