Set Operations in the Unix Shell   v1.01

Created by **Peteris Krumins** (peter@catonmat.net). Peter's blog: **http://www.catonmat.net** – good coders code, great reuse

Set Membership

```
grep -xc 'element' set
grep -xq 'element' set
awk '$0 == "element" { s=1; exit } END { exit !s }' set
awk -v e='element' '$0 == e { s=1; exit } END { exit !s }' set
```

Set Equality

```
diff -q <(sort set1) <(sort set2)
diff -q <(sort set1 | uniq) <(sort set2 | uniq)
awk '{ if (!($0 in a)) c++; a[$0] } END{ exit !(c==NR/2) }' set1 set2
awk '{ a[$0] } END{ exit !(length(a)==NR/2) }' set1 set2
```

Subset Test

```
comm -23 <(sort subset | uniq) <(sort set | uniq) | head -1
awk 'NR==FNR { a[$0]; next } { if !($0 in a) exit 1 }' set subset
```

Set Union   ## Set Intersection

```
cat set1 set2                                   comm -12 <(sort set1) <(sort set2)
awk 1 set1 set2                                 grep -xF -f set1 set2
cat set1 set2 ... setn                          sort set1 set2 | uniq -d
cat set1 set2 | sort -u                         join <(sort -n A) <(sort -n B)
sort set1 set2 | uniq                           awk 'NR==FNR { a[$0]; next } $0 in a' set1 set2
sort -u set1 set2
awk '!a[$0]++'
```

Set Complement   ## Set Cardinality

```
comm -23 <(sort set1) <(sort set2)              wc -l set | cut -d' ' -f1
grep -vxF -f set2 set1                          wc -l < set
sort set2 set2 set1 | uniq -u                   awk 'END { print NR }' set
awk 'NR==FNR { a[$0]; next } !($0 in a)' set2 set1
```

Set Symmetric Difference

```
comm -3 <(sort set1) <(sort set2) | sed 's/\t//g'
comm -3 <(sort set1) <(sort set2) | tr -d '\t'
sort set1 set2 | uniq -u
cat <(grep -vxF -f set1 set2) <(grep -vxF -f set2 set1)
grep -vxF -f set1 set2; grep -vxF -f set2 set1
awk 'NR==FNR { a[$0]; next } $0 in a { delete a[$0]; next } 1; END { for (b in a) print b }' set1 set2
```

Power Set

```
$ p() { [ $# -eq 0 ] && echo || (shift; p "$@") | while read r ; do echo -e "$1 $r\n$r"; done }
$ p `cat set`
```

Set Cartesian Product

```
while read a; do while read b; do echo "$a, $b"; done < set1; done < set2
awk 'NR==FNR { a[$0]; next } { for (i in a) print i, $0 }' set1 set2
```

Disjoint Set Test   ## Empty Set Test

```
comm -12 <(sort set1) <(sort set2)              wc -l set | cut -d' ' -f1
awk '++seen[$0] == 2 { exit 1 }' set1 set2      wc -l < set
                                                awk '{ exit 1 }' set
```

Minimum

```
head -1 <(sort set)
awk 'NR == 1 { min = $0 } $0 < min { min = $0 } END { print min }'
```

Maximum

```
tail -1 <(sort set)
awk 'NR == 1 { max = $0 } $0 > max { max = $0 } END { print max }'
```