Command Line Programs

© MigrateUp, LLC - PowerfulPython.com

A Key To Automation

Much of the power of software comes from its automation.

A good command line interface makes the software you write more automatable. This ALWAYS increases the value of your program.

```
find . -name __pycache__ -type d
ls -l
git commit -m "Fix for #742"
grep -i 'python' *.txt
grep '^WARNING:' logs/*.log
```

© MigrateUp, LLC - PowerfulPython.com

sys.argv

One approach: use `sys.argv`.

```
# findpattern.py
import sys
def grepfile(pattern, path):
    with open(path) as handle:
        for line in handle:
            if pattern in line:
                yield line.rstrip('\n')

if __name__ == "__main__":
    pattern, path = sys.argv[1], sys.argv[2]
    for line in grepfile(pattern, path):
        print(line)
```

```
$ python3 findpattern.py WARNING webserver.log
```

This works. But it makes you do the heavy lifting.

© MigrateUp, LLC - PowerfulPython.com

Optional Arguments

If you have multiple optional arguments, it gets really complex:

- `-i` for case-insensitive search
- `-c` for line count (instead of printing matching lines)
- long (`--count`) and short (`-c`) options that mean the same
- `--limit NUM` to show only `NUM` matches
- conflicting options that should not be used together
- Detect their presence in any order

```
$ python3 findpattern.py -i 'WARNING' webserver.log -c
```

The complexity explodes. We need a better approach.

© MigrateUp, LLC - PowerfulPython.com

sys.argv

Also, the error messages are weird:

```
$ python3 findpattern.py
Traceback (most recent call last):
  File "findpattern.py", line 10, in <module>
    pattern, path = sys.argv[1], sys.argv[2]
IndexError: list index out of range
```

© MigrateUp, LLC - PowerfulPython.com

argparse

The modern tool for building a command line interface is `argparse`.

- Easily specify complex sets of required and optional arguments
- Uses declarative syntax. Python handles the logic for you
- Boolean flags and other type conversions
- Automatic, user-friendly online help
- Clear and understandable error messages

```
import argparse
parser = argparse.ArgumentParser()
```

© MigrateUp, LLC - PowerfulPython.com

Using argparse

```
# findpattern.py, version 2:
# Use the same grepfile() function.

import argparse # instead of "import sys"

parser = argparse.ArgumentParser()
parser.add_argument('pattern')
parser.add_argument('path')

if __name__ == "__main__":
    args = parser.parse_args()
    for line in grepfile(args.pattern, args.path):
        print(line)
```

© MigrateUp, LLC - PowerfulPython.com

Parsing Args

`parse_args()` returns the parsed arguments:

```
>>> from findpattern import parser
>>> args = parser.parse_args(["WARNING", "webserver.log"])
>>> args
Namespace(path='webserver.log', pattern='WARNING')
>>> args.pattern
'WARNING'
>>> args.path
'webserver.log'
```

`parse_args()` operates on `sys.argv` by default:

```
# These two are equivalent.
parser.parse_args()
parser.parse_args(sys.argv[1:])
```

© MigrateUp, LLC - PowerfulPython.com

Helpful Errors

`argparse` includes automatic, user-friendly validation:

```
$ python3 findpattern.py
usage: findpattern.py [-h] pattern path
findpattern.py: error: the following arguments are required: pattern, path
```

And look, there's a -h option!

```
$ python3 findpattern.py  -h
usage: findpattern.py [-h] pattern path

positional arguments:
  pattern
  path

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

© MigrateUp, LLC - PowerfulPython.com

More Help

```
parser = argparse.ArgumentParser(
    description='Find patterns in file',
    epilog='Similar to grep, but with substring matching only.')
parser.add_argument('pattern', help='Substring pattern to match')
parser.add_argument('path', help='File to search in')
```

```
$ python3 findpattern.py -h
usage: findpattern.py [-h] pattern path

Find patterns in file

positional arguments:
  pattern     Substring pattern to match
  path        File to search in

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

Similar to grep, but with substring matching only.
```

© MigrateUp, LLC - PowerfulPython.com

Boolean Flags

```
parser.add_argument('pattern', help='Substring pattern to match')
parser.add_argument('path', help='File to search in')
parser.add_argument('-i', '--ignore-case', action='store_true')
```

```
>>> args = parser.parse_args(["-i", "WARNING", "webserver.log"])
>>> args.ignore_case
True
>>>
>>> args = parser.parse_args(["--ignore-case", "WARNING", "webserver.log"])
>>> args.ignore_case
True
>>>
>>> args = parser.parse_args(["WARNING", "webserver.log"])
>>> args.ignore_case
False
```

You can also invert the logic with `action="store_false"`.

© MigrateUp, LLC - PowerfulPython.com

Option Args

```
parser.add_argument('-p', '--prefix', default='',
                    help='Prepend prefix string to output lines')
```

```
>>> args = parser.parse_args(["--prefix", "FOUND: ", "WARNING",
"webserver.log"])
>>> args.prefix
'FOUND: '

>>> args = parser.parse_args(["WARNING", "webserver.log"])
>>> args.prefix
''
```

© MigrateUp, LLC - PowerfulPython.com

Typed Arguments

By default, arguments are parsed as strings. But you can specify a type.

```
parser.add_argument('--limit', default=None, type=int,
    help='Show only this many matches. Default is show all')
```

```
>>> args = parser.parse_args(["--limit", "42", "WARNING", "webserver.log"])
>>> print(args.limit)
42
>>> type(args.limit)
<class 'int'>

>>> args = parser.parse_args(["WARNING", "webserver.log"])
>>> print(args.limit)
None
```

© MigrateUp, LLC - PowerfulPython.com

Default

Options need a default value, for when they're not supplied.
But if you don't specify `default`, the default value is `None`.

```
# Previous version:
parser.add_argument('--limit', default=None, type=int,
    help='Show only this many matches. Default is show all')

# But you can shorten it to this:
parser.add_argument('--limit', type=int,
    help='Show only this many matches. Default is show all')
```

Exception: if you use `action="store_true"`, the default is `False`.
(And vice versa for `"store_false"`.)

```
# We had written the -i option this way:
parser.add_argument('-i', '--ignore-case', action='store_true')

# Which is just a shortcut for this:
parser.add_argument('-i', '--ignore-case', default=False,
                    action='store_true')
```

© MigrateUp, LLC - PowerfulPython.com

Validation and Type

Using `type` with `add_argument()` will extend `ArgumentParser`'s validation and help output:

```
parser.add_argument('--limit', type=int,
    help='Show only this many matches. Default is show all')
```

```
$ python3 findpattern.py --limit notanumber WARNING webserver.log
usage: findpattern.py [-h] [-i] [--limit LIMIT] pattern path
findpattern.py: error: argument --limit: invalid int value: 'notanumber'
```

Also: the exit code of the process is non-zero.

Basically, `parser.parse_args()` will call `sys.exit(2)` if your program isn't invoked correctly.

© MigrateUp, LLC - PowerfulPython.com

Custom "type"

```
def positive_int(value):
    "Converts string value into a positive integer."
    number = int(value)
    if number <= 0:
        raise ValueError("Bad value: " + str(value))
    return number

# ...

parser.add_argument('--limit', type=positive_int,
    help='Show only this many matches. Default is show all')
```

```
$ python3 findpattern.py --limit 0 WARNING webserver.log
usage: findpattern.py [-h] [-i] [--limit LIMIT] pattern path
findpattern.py: error: argument --limit: invalid positive_int value: '0'
```

© MigrateUp, LLC - PowerfulPython.com

Mutually Exclusive Options

Have a set of options that should not be used together? You can use a *mutually exclusive group*:

```
parser = argparse.ArgumentParser()
quantity_group = parser.add_mutually_exclusive_group()
quantity_group.add_argument('-c', '--count', action='store_true',
                    help='Show # of matches instead of matching lines')
quantity_group.add_argument('--limit', type=positive_int,
                    help='Show only this many matches. Default is show all')
args = parser.parse_args()
```

```
$ python3 findpattern.py --limit 7 -c WARNING webserver.log
usage: findpattern.py [-h] [-c | --limit LIMIT]
findpattern.py: error: argument -c/--count: not allowed with argument --limit
```

© MigrateUp, LLC - PowerfulPython.com

Custom Validation

Alternative: do the check yourself, and use `parser.error()`.

```
parser = argparse.ArgumentParser()
parser.add_argument('-c', '--count', action='store_true',
                    help='Show # of matches instead of matching lines')
parser.add_argument('--limit', type=positive_int,
                    help='Show only this many matches. Default is show all')
args = parser.parse_args()
if args.count and args.limit:
    parser.error('Cannot combine --limit and --count options')
```

```
$ python3 findpattern.py --limit 7 -c WARNING webserver.log
usage: findpattern.py [-h] [-i] [-c] [--limit LIMIT] pattern path
findpattern.py: error: Cannot combine --limit and --count options
```

Downside: Must do more work. Repeating yourself in the help text.
Upside: Potentially more readable.
But: `error()` can handle situations `add_mutually_exclusive_group()` can't.

© MigrateUp, LLC - PowerfulPython.com

Encapsulating

If your parser gets a little complex, encapsulate it in a function.

```
import argparse
def get_args():
    parser = argparse.ArgumentParser()
    parser.add_argument('pattern', help='Substring pattern to match')
    parser.add_argument('path', help='File to search in')
    parser.add_argument('-i', '--ignore-case', action='store_true')
    parser.add_argument('-c', '--count', action='store_true',
                        help='Show # of matches instead of matching lines')
    parser.add_argument('--limit', type=positive_int,
        help='Show only this many matches. Default is show all')
    args = parser.parse_args()
    if args.count and args.limit:
        parser.error('Cannot combine --limit and --count options')
    return args
    # If you don't need to do any special post-parse
    # checking, the last line will simply be:
    # return parser.parse_args()

if __name__ == "__main__":
    args = get_args()
```

© MigrateUp, LLC - PowerfulPython.com

Better UX

```
if __name__ == "__main__":
    args = get_args()
```

For a better user experience, parse the args immediately inside the main block.

Quicker feedback from `--help`, error checking, etc.

© MigrateUp, LLC - PowerfulPython.com

Lab: Command Line Arguments

Lab file: `commandline/argparselab.py`

Reference:

https://docs.python.org/3/library/argparse.html

(Linked as "Argparse Reference" under this video. There's also a link in top of your lab file, for easy copy-pasting.)