

Report on the python hashing implementation.

General Overview:

pyHasher normalizes a file and creates a hash for it. The normalization is done by running mnfy3 or mnfy2, and if both fail, falling back to mnfyfailsafe, then creating a sha512 hash of the output file. It can be run in either python2 or python3, as long as both are installed.

Usage instructions:

To hash a single file, use the following command:

```
'pyHasher.py <filename>'
```

Optionally, to see the normalized file, append '—dump'. (though this is only useful for debugging/test)

Assumptions:

Differences between ASTs:

Python 2.4	Python 2.5	Python 2.6	Python 3.3
*Cant be found/Does not exist	ImportFrom module was not optional in 2.5 Set/SetComp/DictComp were introduced in 2.7	ImportFrom module was not optional in 2.5 Set/SetComp/DictComp were introduced in 2.7	

Changes Made and Status:

Change	Developed	Tested	Notes

Known Issues:

Issue	Location	Impact	Severity	Action
Slicing error- P=(p+[1,2,3]):1] Gets minified to P=p+[1,2,3]:1] Which has different meaning.	Mnfy27, Mnfy33	No impact on hashing because its consistant. Is an issue for mnfy's standalone goals of portable code.	Low	No action required for hashing /RedHat purposes.

Some docstrings are referenced by <code>__doc__</code> call in python. Docstrings are removed by mnfy, causing compile/run problems.	Mnfy27, Mnfy33	No impact on hashing because code does not need to compile and removal of docstrings is consistent.	Low	No action required for hashing/RedHat purposes.
Big Numbers: For some reason, reading and printing very large numbers (in the order of 10^{10}) causes digits to be lost.	Mnfy27, but I think it is in all of Python.	Potential hashing risk, because large numbers could change code execution, though both hashes would be the same.	Medium	Not really sure if there is a fix for this one.

Testing Report:

Test scripts:

/libs/mnfy33/test_mnfy.py - Unit Testing of Mnfy33 (written by Brett Cannon)

/libs/mnfy33/verify_mnfy.py - Coverage Testing (written by Brett Cannon)

/libs/mnfy273/test_mnfy.py - Unit Testing of Mnfy27 (modified from mnfy3300 version)

/libs/mnfy273/verify_mnfy.py - Coverage Testing of Mnfy27 (modified from mnfy3300 version)

Unit Testing of MnfyFailSafe – Informally tested, no test scripts saved.

testHasher.py – Testing of hashing function/blackbox testing/testing multiple environments.

*Note: Not all tests pass, because of the issues listed above.