

# Testausdokumentti

Testaus toteutaan yksikkötesteillä ja kokonaisvaltaisina pakkaustesteinä. Yksikkötestit tarkastavat jokaisen pienen funktion toiminnan, ja kokonaisvaltaiset testit koko ohjelman toiminnan.

Testitiedostoina käytettynä on:

- a\_only: Yli 5 MB pelkkää pientä a-kirjainta.
- declaration\_inde: Yhdysvaltojen itsenäisyydenjulistus
- anti\_oedipus: Pitkä pätkä Gilles Deleuzen ja Félix Guattarin kirjasta Anti-Oidipus
- random: Satunnaisesti generoitu pitkä tekstipätkä
- one: Yksi b-kirjain
- large: 100 MB satunnainen merkkijono
- zero\_same: Merkkijono, jossa yksikään merkki ei esiinny useammin kuin kerran

Yleiset pakkaustestit suoritetaan molemmille algoritmeille niin, että jokainen kyseisistä tiedostoista pakataan ja puretaan uudelleen. Pakattua kokoa verrataan alkuperäiseen kokoon, ja katsotaan kuinka pieneen tilaan se saatiin. Tämän jälkeen tarkistetaan, että uudelleenpurettu ja alkuperäinen tiedosto ovat samat. Algoritmien suorituskyykyä ei testata yksinkertaisesti siitä syystä, että python-kielen luonteen mukaan suorituskyykyerät eivät ole kovinkaan merkittäviä. Kumpikaan algoritmi ei ole salamannopea. Etenkin suurten tiedostojen, kuten 100MB-tiedoston pakkaus vie aikaa. Testitiedostot sijaitsevat test\_files/ -kansiossa.

Testaus aloitetaan ensin asentamalla poetry

**poetry install**

**poetry shell**

**poetry tests/huffman\_test\_compression.py**

Tässä Huffman-koodauksen testin tulokset. Huffman-koodaus onnistuu purkamaan jokaisen tiedoston, ja ainoa pakattu tiedosto, joka jää suuremmaksi kuin alkuperäinen on ei ollenkaan toistuvuutta sisältävä tiedosto. Koodaus onnistuu jopa pakkaamaan 100MB tiedoston hyvin.

```
(algoai-py3.13) PS C:\Users\Eetos\algoai> python tests/huffman_test_compression.py
File: anti_oedipus.txt
Original: 1183836 bytes
Compressed: 651878 bytes
Decompressed:1183836 bytes
Original and decompressed match: True
Ratio: 0.5506489074500184
File: a_only.txt
Original: 5829200 bytes
Compressed: 728694 bytes
Decompressed:5829200 bytes
Original and decompressed match: True
Ratio: 0.12500754820558568
File: declaration_inde.txt
Original: 8213 bytes
Compressed: 5456 bytes
Decompressed:8213 bytes
Original and decompressed match: True
Ratio: 0.6643126750273955
File: large.txt
Original: 104857600 bytes
Compressed: 13107251 bytes
Decompressed:104857600 bytes
Original and decompressed match: True
Ratio: 0.12500048637390138
File: one.txt
Original: 1 bytes
Compressed: 39 bytes
Decompressed:1 bytes
Original and decompressed match: True
Ratio: 39.0
File: random.txt
Original: 339998 bytes
Compressed: 239139 bytes
Decompressed:339998 bytes
Original and decompressed match: True
Ratio: 0.7033541373772787
File: zero_same.txt
Original: 33 bytes
Compressed: 411 bytes
Decompressed:33 bytes
Original and decompressed match: True
Ratio: 12.454545454545455
(algoai-py3.13) PS C:\Users\Eetos\algoai> |
```

### poetry tests/lz78\_test\_compression.py

Vastaavasti taas LZ78-algoritmi saa erilaisia tuloksia. Nyt 100MB tiedostoa ei pystytä pakata puhtaasti suuren koon vuoksi. Runsaasti toistuvuutta sisältävät tiedostot saadaan pakattua erittäin pieneen tilaan.

```
(algoai-py3.13) PS C:\Users\Eetos\algoai> python tests/lz78_test_compression.py
File: anti_oedipus.txt
Original: 1183836 bytes
Compressed: 695650 bytes
Decompressed:1183836 bytes
Original and decompressed match: True
Ratio: 0.5876236235424501
File: a_only.txt
Original: 5829200 bytes
Compressed: 8537 bytes
Decompressed:5829200 bytes
Original and decompressed match: True
Ratio: 0.001464523433747341
File: declaration_inde.txt
Original: 8213 bytes
Compressed: 5700 bytes
Decompressed:8213 bytes
Original and decompressed match: True
Ratio: 0.6940216729575064
File: large.txt
Original: 104857600 bytes
Compressed: 39828 bytes
Decompressed:0 bytes
Original and decompressed match: False
Ratio: 0.0003798294067382812
File: one.txt
Original: 1 bytes
Compressed: 4 bytes
Decompressed:1 bytes
Original and decompressed match: True
Ratio: 4.0
File: random.txt
Original: 339998 bytes
Compressed: 318359 bytes
Decompressed:339998 bytes
Original and decompressed match: True
Ratio: 0.9363555079735764
File: zero_same.txt
Original: 33 bytes
Compressed: 60 bytes
Decompressed:33 bytes
Original and decompressed match: True
Ratio: 1.8181818181818181
(algoai-py3.13) PS C:\Users\Eetos\algoai> |
```

Yksikkötestit voidaan suorittaa komenolla:

**pytest -v**

Jokainen yksikkötesti menee läpi.

```
(algoai-py3.13) PS C:\Users\Eetos\algoai> pytest -v
===== test session starts =====
platform win32 -- Python 3.13.7, pytest-8.4.2, pluggy-1.6.0 -- C:\Users\Eetos\AppData\Local\pypoetry\Cache\virtualenvs\algoai-qtskMdUd-py3.13\Scripts\python.exe
cachedir: .pytest_cache
rootdir: C:\Users\Eetos\algoai
configfile: pyproject.toml
collected 8 items

tests/huffman_test.py::TestHuffmanCodec::test_dictionary_and_tree_creation PASSED [ 12%]
tests/huffman_test.py::TestHuffmanCodec::test_encoding_and_decoding PASSED [ 25%]
tests/huffman_test.py::TestHuffmanCodec::test_missing_file_raises PASSED [ 37%]
tests/huffman_test.py::TestHuffmanCodec::test_roundtrip_file PASSED [ 50%]
tests/lz78_test.py::TestLZ78Codec::test_encoding_and_decoding_match PASSED [ 62%]
tests/lz78_test.py::TestLZ78Codec::test_nonexistent_file_raises PASSED [ 75%]
tests/lz78_test.py::TestLZ78Codec::test_roundtrip_file PASSED [ 87%]
tests/lz78_test.py::TestLZ78Codec::test_save_and_load_compressed PASSED [100%]

===== 8 passed in 0.73s =====
(algoai-py3.13) PS C:\Users\Eetos\algoai> |
```

Testikattavuusraportin saa tulostettua komennolla:

**coverage report -m**

```
(algoai-py3.13) PS C:\Users\Eetos\algoai> coverage report -m
Name                               Stmts   Miss Branch BrPart  Cover   Missing
-----
algoai\huffman\codec.py            109      4      24      0    95%    16-19, 71
algoai\lz78\codec.py                87      3      28      3    95%      6, 46, 79
tests\huffman_test.py              48      1       2      1    96%      59
tests\lz78_test.py                 43      1       2      1    96%      53
TOTAL                             287      9      56      5    95%
(algoai-py3.13) PS C:\Users\Eetos\algoai> |
```

Raportti antaa hyvät tulokset.