|  |  |
| --- | --- |
|  |  |
| LZ77 Compression Technique |  |
|  |  |
|  | DSAI 325Information Theory |
|  | Mohamed Ehab Yousri202201236 |

LZ77 Compression Technique

is a lossless data compression method. It works by replacing repeated sequences in the text with references to their previous occurrences, effectively reducing the overall size of the data.

Implementation Details:

Compression Algorithm:

The compression processes are carried out in the following manner:

-Keeps a search buffer where the text that has been previously encountered is stored.

-Look through the text with each character to find the longest possible match in the search -buffer.

-Records the match as a triplet (Offset, Length, Next Character).

-Keeps scanning until the input text is completely compressed.

Decompression Algorithm:

The decompression processes are as follows:

-Comprehends the compressed triplets (Offset, Length, Next Character).

-Recovers the original paragraph by looking up previously decoded characters and using them as references.

-Adds the required next character to meet the text completion.

The Main Components:

**Search Buffer** → Stores already seen characters.

**Lookahead Buffer** → Contains upcoming characters to be processed.

The Algorithm Implementation:

Compression Algorithm:

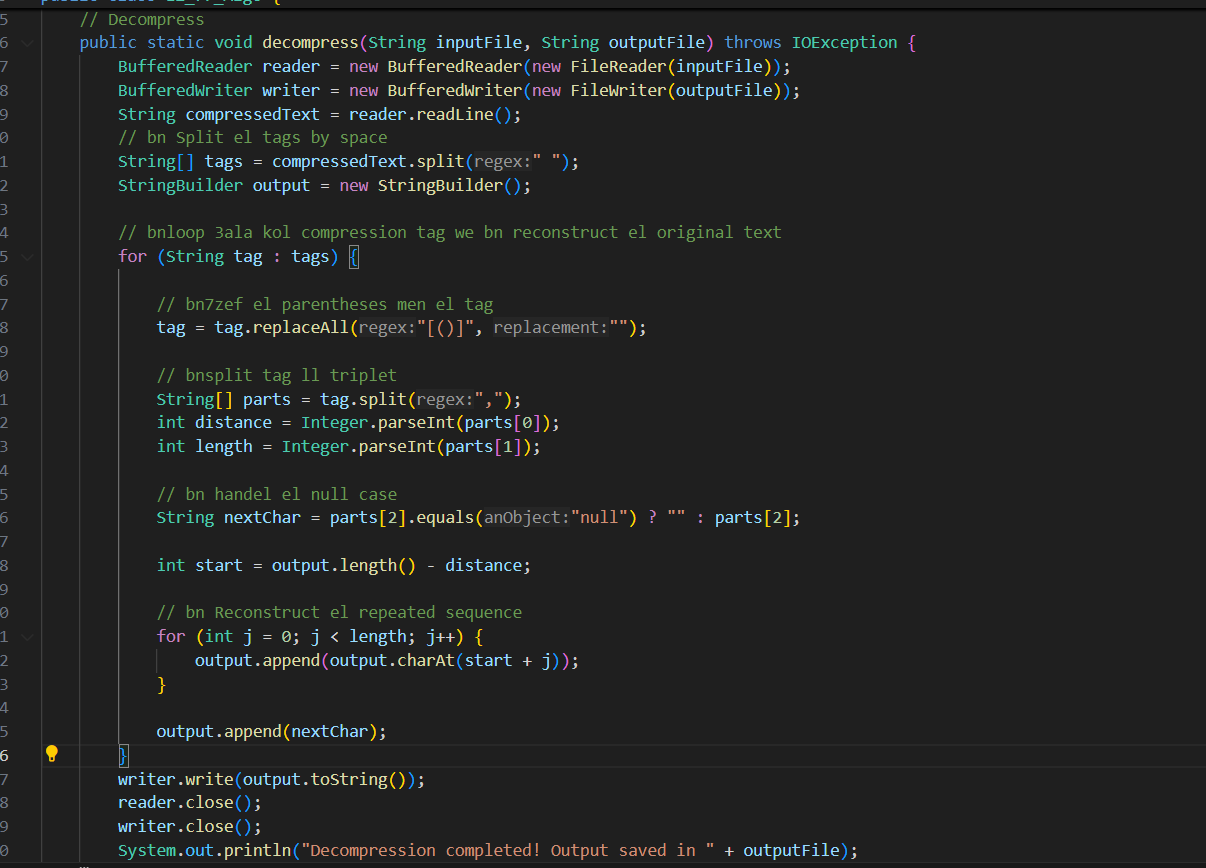
A screen shot of a computer program

AI-generated content may be incorrect.

A screen shot of a computer code

AI-generated content may be incorrect.

Decompression Algorithm:



Main :



Test Cases:

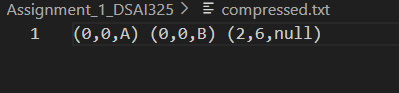
Test Case 1: Sequence:

Original:



**Original Size:** 8 characters × 8 bits = 64 bits

Compressed:



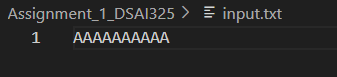
**Compressed Size:** 3 triplets × (2+2+8) bits = 36 bits

**Compression Ratio:** (36 / 64) × 100 = 56.25%

repeated sequence is handled effectively

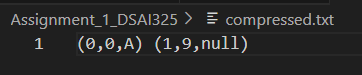
Test Case 2: Repeating Characters

Original:



**Original Size:** 10 characters × 8 bits = 80 bits

Compressed:



**Compressed Size:** 2 triplets × (2+2+8) bits = 24 bits

**Compression Ratio:** (24 / 80) × 100 = 30%

repeated sequence is handled effectively

Test Case 3: Long Repetitive Sequence

Original:

A screenshot of a computer

AI-generated content may be incorrect.

**Original Size:** 16 characters × 8 bits = 128 bits

Compressed:

A screenshot of a computer

AI-generated content may be incorrect.

**Compressed Size:** 3 triplets × (2+2+8) bits = 36 bits

**Compression Ratio:** (36 / 128) × 100 = 28.12%

repeated sequence is handled effectively

The Full Running Process :

A screenshot of a computer

AI-generated content may be incorrect.

References

<https://www.youtube.com/watch?v=PrbTpYVDv6Q>

<https://www.youtube.com/watch?v=goOa3DGezUA&t=40s>

<https://medium.com/@vincentcorbee/lz77-compression-in-javascript-cd2583d2a8bd>

<https://learn.microsoft.com/en-us/openspecs/windows_protocols/ms-wusp/fb98aa28-5cd7-407f-8869-a6cef1ff1ccb>

https://www.youtube.com/watch?v=zev2A2uWYsM