

# LZ77-komprimeringsfamilien

Adam og Suleyman

# Oversikt

- LZ77
- DEFLATE
- ZSTD

# Introduksjon til LZ77

- Designet av Abraham Lempel og Jacob Ziv (1977)
- Tapsfri datakomprimeringsalgoritme
- Type: Dictionary coder

# Sliding window

- Kjernekonsept i LZ77
- Fungerer som en glidende ordbok
- Brukes til å finne tidligere mønstre

# Ordbok Oppbygging

- Deler opp datasekvensen i unike deler

*AABABBBABAABABBBABBABB*

- Tillater duplikater (f.eks. "BB")

*A|AB|ABB|B|ABA|ABAB|BB|ABBA|BB*

- Referanse: MIT Math Department

# Tabellstruktur

- Tre hovedkomponenter:

- Indeks til ordet	1	2	3	4	5	6	7	8	9
- Selve ordet	<i>A</i>	<i>AB</i>	<i>ABB</i>	<i>B</i>	<i>ABA</i>	<i>ABAB</i>	<i>BB</i>	<i>ABBA</i>	<i>BB</i>
- Kode (oppskrift)	$\emptyset A$	<i>1B</i>	<i>2B</i>	$\emptyset B$	<i>2A</i>	<i>5B</i>	<i>4B</i>	<i>3A</i>	<i>7</i>

- Kan bruker -1 som markør for tomt sett i implementasjonen

- Koden består av indeks + suffix

- Brukes til å bygge opp ordet
- Bruker indeks til å finne neste prefix til suffixen
- *ABBA* -> *3A*, *2B*, *1B*,  $\emptyset A$  -> *ABBA* (Husk at den må reversert)
- Alle forekomst av *ABBA* kan derfor bli erstattet med *3A*
- *FOOABBAR* -> *FOO3AR*

- Trie eller ordbok

# Hovedprinsippene bak LZ77

- Sliding window-teknikk
- Bygger ordbok
- Erstatter duplikater i lookahead buffer
- Avhengig på implementasjonen

# Deflate

- Designa av Phil Katz
- Brukes i zip og gzip filer
- LZ77 og Huffman koding
- Block header
  - **00** ikke komprimert
  - **01** komprimert med fixed huffman koding
  - **10** komprimer med dynamisk huffman koding
  - **11** reservert



# ZSTD

- Utviklet av Yann Collett
- LZ77 og Huffman koding
- FSE(finite-state entropy)
  - Tilstand

# ZSTD

- Komprimeringsnivå
  - lav nivå-> raskere komprimering/dekomprimering, dårligere komprimeringsforhold
  - høy nivå -> treigere komprimering/dekomprimering, større komprimeringsforhold
- Training mode
  - Dictionary

# Kildehenvisning

Lossless Data Compression: LZ78. (n.d.). Cs.stanford.edu.

<https://cs.stanford.edu/people/eroberts/courses/soco/projects/data-compression/lossless/lz78/concept.htm>

Peter Shor (14 October 2005). "Lempel–Ziv notes"

[https://web.archive.org/web/20210528171521/http://www-math.mit.edu/~shor/PAM/lempel\\_ziv\\_notes.pdf](https://web.archive.org/web/20210528171521/http://www-math.mit.edu/~shor/PAM/lempel_ziv_notes.pdf). Archived from the original [https://math.mit.edu/~shor/PAM/lempel\\_ziv\\_notes.pdf](https://math.mit.edu/~shor/PAM/lempel_ziv_notes.pdf) on 28 May 2021. Retrieved 9 November 2014.

Deutsch, P. (1996). *DEFLATE compressed data format specification version 1.3* (No. rfc1951).

Harnik, D., Khaitzin, E., Sotnikov, D., & Taharlev, S. (2014, March). A fast implementation of deflate. In *2014 Data Compression Conference* (pp. 223-232). IEEE.

Collet, Y., & Kucherawy, M. (2018). *Zstandard Compression and the application/zstd Media Type* (No. rfc8478).

Facebook. (2024) Zstandard- Zstd. Github <https://github.com/facebook/zstd>

Collet, Y. (2014). Finite State Entropy - A new breed of entropy coder. Blogspot.com.

<https://fastcompression.blogspot.com/2013/12/finite-state-entropy-new-breed-of.html>