

Huffmanovo kódování

je příkladem bezztrátového kompresního algoritmu. Nejčastnějším znakům přiřazuje kratší kódovací sekvenci, čímž šetří místo. Je detailně popsané např. na [anglické wiki](#), existuje [online applet](#), který provede rozbor na strom libovolné zadané věty.

Příklad komprese textu

Mějme větu: **TENTO TEXT JE JEN TEST**

V klasickém kódování ASCII zabere 22 B neboli 176 b.

V Huffmanově kódování pouze 46 b tedy necelých 6 B.

Postup:

1. **Zjistíme četnost** jednotlivých znaků:

| | | | | | | | | |
|----------------|---|---|-----|---|---|---|---|---|
| Znak | T | E | SPC | N | J | O | X | S |
| Četnost | 6 | 5 | 4 | 2 | 2 | 1 | 1 | 1 |

2. **Vytvoříme strom** četností: Vezmeme dva znaky s nejnižší četností, vytvoříme z nich nový „složený znak (dvojznak)“, jehož četnost odpovídá součtu četností znaků ve dvojici. To opakujeme tak dlouho, dokud nezbude jen jediný „složený znak“ ze všech znaků obsažených ve zprávě.
3. Zpětně **procházíme vzniklý strom** a na každé úrovni **přidělujeme vždy symboly 1 a 0** tak, že větev/znak s vyšší četností dostane 1 a větev/znak s nižší četností 0.

Na příkladu:

1. Nejnižší četnost mají znaky X,S, proto vznikne dvojznak (X,S) s četností 2:

| | | | | | | | |
|----------------|---|---|-----|---|---|-------|---|
| Znak | T | E | SPC | N | J | (X,S) | O |
| Četnost | 6 | 5 | 4 | 2 | 2 | 2 | 1 |

2. Další dvojice v pořadí je O a dvojznak (X,S):

| | | | |
|-------------|---|---|-----|
| Znak | T | E | SPC |
|-------------|---|---|-----|

From:

<https://old.gml.cz/wiki/> - GMLWiki

Permanent link:

<https://old.gml.cz/wiki/doku.php/informatika:maturita:huffmann?rev=1601153862>

Last update: 26. 09. 2020, 22:57

