

907 : Algorithmique du texte. Exemples et applications.

Antoine DEQUAY

21 septembre 2022

Notes

- Prof : .
- Références :
 - CHRETIENNE,
 - CORMEN,
 - CROCHEMORE, *algorithmique du texte*,
 - INCERTI, *Compression d'image : algorithmes et standards*,
 - SCHWARZENTRUBER.

Table des matières

1	Recherche de motif	1
1.1	Approche naïve	1
1.2	Vers des algorithmes plus efficaces	1
2	Comparaison de mots	1
2.1	Plus longue sous-séquence commune	1
2.2	Distance entre 2 mots	1
3	Analyse lexicale et syntaxique	1
3.1	Analyse lexicale	1
3.2	Analyse syntaxique	1

4	Compression	2
4.1	Codage <i>LZ77</i>	2
4.2	Codage de HUFFMAN	2
4.3	Codage <i>LZ78</i>	2

1 Recherche de motif

Cf CHRETIENNE *et* CORMEN.

1.1 Approche naïve

↪ Def alphabets, mot, préfixe, ...

↪ Algo naïf.

1.2 Vers des algorithmes plus efficaces

↪ Automate des occurrences avec KM,

↪ [DEV] KMP,

↪ Remarque : on peut adapter à la recherche d'expressions rationnelles, voir plus loin avec analyse lexicale.

2 Comparaison de mots

2.1 Plus longue sous-séquence commune

Cf CROCHEMORE *ou* CORMEN.

2.2 Distance entre 2 mots

Cf CROCHEMORE.

3 Analyse lexicale et syntaxique

Cf SCHWARZENTRUBER.

3.1 Analyse lexicale

3.2 Analyse syntaxique

↪ [DEV] CYK,

↪ Si place : ([DEV]) Earley.

4 Compression

Cf INCERTI.

4.1 Codage LZ77

↪ Pour culture : LZSS améliore l'algo en utilisant un bit au lieu de 3 octets pour les nouveaux caractères et en utilisant un arbre pour stocker le dictionnaire.

4.2 Codage de HUFFMAN

4.3 Codage LZ78

↪ Codage LZ78,

↪ Codage LZW.